Visualisation and Exploration of Personal Data in Virtual Reality

Patrick Millais

Bachelor of Science in Computer Science with Honours
The University of Bath
May 2017

Visualisation and Exploration of Personal Data in Virtual Reality

Submitted by: Patrick Millais

COPYRIGHT

Declaration

This dissertation is submitted to the University of Bath in accordance with the requirements of the degree of Bachelor of Science in the Department of Computer Science. No portion of the work in this dissertation has been submitted in support of an application for any other degree or qualification of this or any other university or institution of learning. Except where specifically acknowledged, it is the work of the author.

Signed:

Abstract

Recent research in the personal informatics field has focused on correlating aspects of self-tracked data, supporting users to arrive at meaningful insights when reflecting on aggregated datasets. To date, no research has been completed on how users could explore personal data using virtual reality, and the opportunities this presents for users' understanding of multidimensional datasets.

In this study we evaluate the open-ended exploration of multidimensional datasets using two separate visualisations. *Be The Data* immerses users in a three-dimensional scatter plot, allowing them to interpret a dataset from new perspectives. The second visualisation, *Parallel Planes*, enables a multi-faceted dataset to be chained together, supporting users in perceiving a holistic overview of interrelated dimensions.

Through an insight-based evaluation methodology, we find that users conducted depth-based explorations of the Parallel Planes visualisation, arriving at valuable and significant insights through hypothesising about the data. We also find that there was no overall task-workload difference between traditional visualisation paradigms and virtual reality. We conclude by outlining future research directions, and making recommendations for future evaluation approaches for data visualisation in VR.

Contents

List of Figures

List of Tables

Acknowledgements

I have been very lucky to have such a brilliant supervisor who has guided me with thoughtful comments for the past year. Simon, thank you.

In the spirit of self-tracking, thanks to the 2,996 artists and 17,055 songs which have provided the soundtrack to this dissertation – it's been an incredible year for music.

Lastly, many thanks also to my family for their (financial) support, and friends and flatmates for making final year so enjoyable.

Chapter 1

Introduction

The proliferation of personal computers, connecting and interacting through the Internet, is the defining characteristic of the Information Age. Entire economies are now built around data collected from personal devices, revolutionising modern approaches to healthcare, financial services and education. *Prosumption*, first termed by Toffler (1981), describes the intersection in roles between producers and consumers. The rise of ubiquitous computing, driven by prosumer-generated content, presents a significant challenge for the data visualisation field. What challenges do non-expert users face when viewing and consuming their personal data? How can people draw meaningful insights from visualisations of their own data?

Personal analytics platforms are now beginning to emerge, supporting an extensive number of data tracking services such as Fitbit, Last.FM and various social media platforms. Exist.io is one example of an aggregation platform, identifying data correlations between separate services to provide habit insights in the form of graphical visualisations presented on data dashboards. These dashboards transform distinct data services into aggregated, consumable information providing a *"unique solution to the problem of information overload"* between multi-faceted information (Few, 2013, p.37). Encouraging users to make behavioural changes based on visualisation insights is a current area of research (Whooley et al., 2014). How can data visualisation better stimulate self-reflection in order for users to understand and benefit from their personal data?

One approach may be through Virtual Reality (VR). VR is not a new phenomenon, having been popularised in the early 1990s by companies such as VPL research and VR gaming machines produced by Virtuality Group (Lanier and Biocca, 1992). VR never took off then for consumers, with the effects of nausea, weak hardware and the lack of standards cited as reasons for its early decline (Horowitz, 2004; Arthur, 2015). However, a resurgence is happening. Technological advances induced by Moore's Law provide a platform for consumer-led VR experiences. Stand-alone head-mounted displays such as the Oculus Rift and HTC Vive, alongside low-cost VR experiences provided by Google Daydream in which consumers' smartphones drive the experience, have all emerged over the past 12 months.

Notably with smartphone VR, a single device could be used for both the collection of personal data and reflecting upon this data. Analysts predict extraordinary growth for VR, with the addressable market size in 2025 larger than today's television market (Bellini et al., 2016).

Bryson (1996, p.62) defines *virtual reality* as *"the use of computers and human-computer interfaces to create the effect of a three-dimensional world containing interactive objects with a strong sense of three-dimensional presence"*. Establishing *presence* is a key area of research for developers. Maximising sensory input through three-dimensional audio and haptic feedback helps to build a sense of presence in VR (Van Dam et al., 2000). Indeed, data visualists look towards VR's wider field of view to provide global context for scientific data, enabling more natural and quicker exploration of big data sets (Van Dam et al., 2000). Such active, participatory interaction techniques in VR are therefore a fundamental area of interest for this project, given the ability of VR to engage users with the content and promote further data exploration.

Existing personal informatics systems can support increased self-understanding by revealing relationships between wellbeing aspects (e.g food intake and sleep) (Bentley et al., 2013). However, current visualisation of statistical patterns are presented through natural language, or statistical graphics, with no research around visualisations of personal data in virtual reality environments. This project will consider the opportunities of visualising personal data in VR and the challenges this brings for understanding and interacting with multi-dimensional personal datasets.

1.1 Aims and Contributions

The aim of this project is to explore the overlap of personal informatics, visualisation, and virtual reality fields. To the best of our knowledge, there is no existing work in this area, and as a result our study is rather exploratory in nature. Our objective is therefore to investigate the differences between personal data exploration in traditional paradigms and in VR, and the opportunities and challenges this brings for data reflection.

We develop two different techniques of visualising personal data with varying dimensionality. Specific contributions of this project include:

- *Be The Data:* A three-dimensional scatter graph represented in a virtual world. Users can navigate around in this environment and take on the perspective of a data point, immersing themselves in the personal dataset, and interpret the visualisation from new angles.
- *Parallel Planes:* An extension of a parallel coordinates visualisation, adding a dimension in the z axis. The visualisation enables a multi-faceted dataset to be chained together, letting the user perceive a holistic overview of interrelated dimensions.
- An extensive empirical evaluation measuring presence, task workload and insight for

both of the visualisations. Significantly, we suggest that presence may not be the most appropriate evaluation method for assessing data visualisations in VR due to its close ties to realism. We also make recommendations for future studies looking to use Saraiya et al. (2005)'s insight-based evaluation approach, and discuss future research directions.

1.2 Outline

The structure of this dissertation is as follows:

Chapter 1 outlines the project and introduces the motivation behind this work.

Chapter 2 explores three fields in terms of the literature: personal informatics, visualisation, and virtual reality. We identify the two visualisation techniques which this project evaluates: *Be The Data*, *Parallel Planes*.

Chapter 3 describes the requirement elicitation process. A non-exhaustive list of requirements are produced, which are later refined and scoped during the prototyping process.

Chapter 4 focuses on the key design considerations made during the prototyping process.

Chapter 5 displays the implementation results of our visualisations in a graphical format.

Chapter 6 revisits the literature to justify our chosen evaluation method. The experimental design is then detailed and hypotheses are formalised.

Chapter 7 discusses and analyses the results of our three-fold evaluation approach, measuring insight, presence and task workload.

Chapter 8 continues the discussion of results in relation to research questions, summarises our contribution and limitations, and sets out future research directions.

Chapter 2

Literature Review

In this chapter, we explore three fields: personal informatics, visualisation, and virtual reality. We begin by investigating the rise of personal data collection, and its association with the stage-based model of informatics. We identify motivations for self-tracking, and the challenges which users face when they reflect upon their collected data.

Various approaches towards communicating personal data through visualisation are examined. In particular, we look at different methods of visualising high-dimensionality datasets. We then establish two separate visualisation techniques: *Parallel Planes* and *Be The Data*. Finally, we describe the relation of presence and immersion, enabling us to evaluate the potential of visualising personal datasets in virtual reality.

2.1 Personal Data

The Quantified Self[1] movement has rapidly gained momentum over the last few years with the arrival of mass market products dominating the wearable technology space. Coupled with the widespread adoption of mobile phones, these devices represent a new era in data production and consumption through smart, always-on sensors. Researchers have explored the benefits of self-tracking, both in terms of understanding the patterns between activities and emotions (Rachuri et al., 2010), and as a means to promote behavioural change (Bentley et al., 2013). Therefore, quantified self is concerned with the collection of data relating to oneself. This is *personal data*, broadly defined by public bodies like the Information Commissioner's Office and the European Commission as any data which relates to a living individual who can be identified from those data and/or other information in possession of the data controller. Indeed, *sensitive* personal data extends to information which could be used in a discriminatory manner - e.g ethnicity, race, political opinions and, in particular, physical and mental health conditions.

However, given that the driving force of the quantified self movement has been health

[1] http://quantifiedself.com/

tracking, we are seeing the emergence of a new classification of self-tracked personal data. This encompasses additional physiological data such as heart rate, body temperature and blood glucose levels. It is an increasingly dynamic term compared to the relatively static IC and EC definitions of personal data. Nevertheless, an EU advisory body has clarified the scope of personal data to include a broader view on health data in new regulations applicable from 2018 (European Commission, 2016). Self-tracked personal data falling under this classification is therefore of particular interest towards this project.

2.1.1 The Rise Of Personal Data Tracking And Analytics

Experimental physiology dates back to the 16th century in one of the earliest examples of quantified self-tracking. Santorio Santorio measured his body weight, food intake and excrement for 3 decades, investigating the characteristics of metabolism (Neuringer, 1981). Fast forward 400 years and self-tracking devices are now rife. Mobile phones are ubiquitous in developed countries and wearable technology has made significant traction across these markets. Research firm Gartner predicts that global shipments of wearable devices will exceed the 500 million mark in 2020 (Levy, 2015). Emerging markets such as the Middle East, Africa and the Asia Pacific regions are bringing an increasing number of smartphones online each year (Ericsson, 2015), opening extraordinary opportunities for consumers and companies alike.

An expansive list of self-tracking products have flooded the market in developed countries, with a select number exploiting these opportunities correctly for commercial success. The Fitbit[2] product line is an example of an activity tracker accumulating data around a person's fitness activity. Early Fitbit trackers clipped to the person's clothing and offered metrics - distance travelled, sleep tracking and calories burned – on top of a traditional pedometer function. The modern Fitbit product line has evolved into wearable wristbands, leveraging additional sensors to track heart rate and floors climbed.

Sports clothing brand Nike have also delved into the wearable fitness market. One of their first tracking products was the Nike+ standalone sensor, a small sensor which slipped into the sole of a sports shoe. A wireless network is used to transmit the data from the sensor to a companion Nike application on iOS and Android mobile devices. In a trend still widely employed by other fitness companies like Fitbit and Garmin, the app integrates with social networks, allowing fitness activity to be shared and gamified within a local community.

Moore's Law has driven the miniaturisation and cost of modern sensors down, enabling their inclusion in mass market smartwatches and smartphones. Sensors such as accelerometers, gyroscopes and barometers are built into devices, with general purpose APIs allowing developers access to measurements relevant to the applications they are building. For example, consider a car insurance company monitoring the driving safety of its drivers. Bespoke, dedicated hardware would have previously been used to calculate driving characteristics but research has shown that inexpensive, accessible smartphones can now be used as a viable

[2]https://www.fitbit.com/uk

alternative (Johnson and Trivedi, 2011). This general purpose approach led by ubiquitous consumer devices has its merits. Smartwatches such as the Apple Watch have an NFC chip for contactless payments, GPS for location, accelerometer for sleep tracking and a heart rate sensor. This level of sensor diversity appeals to a wide market as consumers no longer have to purchase individual dedicated devices for each use case. Arguably this leads to demographics exploring domains previously inaccessible to them, allowing sensors to collect and quantify information about tracked activities.

A Framework For Personal Informatics

Quantified self, personal analytics and personal informatics are a *"class of systems and practices that help people collect and reflect on personal information"* (Choe et al., 2014, p.1144). Therefore it is a broad term that encompasses many types of tracking, not just limited to fitness and health. Li et al. (2010) introduce a stage-based model of the entire personal informatics system formed of 5 stages: preparation, collection, integration, reflection, and action. Each of these stages can be described as either user-driven or system-driven.

Within the basis of the collection stage, Fitbit and Nike are examples of system-driven pesonal data collection. The wearable products assume the responsibility of continuously recording data. Further examples of system-driven data collection outside of the health and fitness space include Last.fm[3] and RescueTime[4]. Last.fm integrates with multiple music services such as Spotify[5] and iTunes[6], logs the songs which the user listens to and builds up a profile around their music taste. Rescuetime reports time spent in applications and on websites, allowing users to set up goals in order to improve work productivity. The common theme across both instances is background software continously monitoring user activity without input by the user themselves.

Conversely, in a user-driven collection stage, users take an active role in manually inputting activities. The social networking app Untappd[7] has been built around this premise, enabling people to record and rate the beers they drink between pubs. This manual process is also employed by Quantified Selfers with more basic methods of tracking – such as using a spreadsheet for data collection, or even through pen and paper (Choe et al., 2014).

Li et al. (2010) describe the opportunities which separate user-driven and system-driven approaches bring, suggesting that developers and designers should select an approach relevant to the overall system. In the context of more "*extreme*" quantified self-trackers, Choe et al. (2014, p.1151) envision a "*balance between fully automated sensing and manual self-report*". System-driven data collection is somewhat limited by its inability to record specific human traits such as mood or pain, and automatic collection can lead to disengagement with the activity data. Accordingly a balance between the two approaches must

[3]https://www.last.fm
[4]https://www.rescuetime.com
[5]https://www.spotify.com
[6]https://www.apple.com/itunes
[7]https://untappd.com

be met, demonstrating the importance of a flexible, user-centered design when targeting data collection.

2.1.2 What Motivates People To Collect Personal Data?

Understanding *how* people collect data is crucial for building adaptable solutions, but understanding *why* people collect data will lead to better data insights. So far we have discussed the second stage of the 5-stage model. However, the first stage - the preparation stage - is perhaps the biggest barrier to widespread consumer adoption of personal tracking. What motivates people towards looking at personal data tracking systems? At what point do they commit to tracking themselves?

Rooksby et al. (2014) identify five overlapping styles in which tracking is initiated:

- Directive Tracking
- Documentary Tracking
- Collecting Rewards
- Fetishised Tracking
- Diagnostic Tracking

While the research acknowledges that the study cohort does not extend to a general population, these definitions provide a useful starting framework for considering and categorising the various factors of motivation. Significantly these are styles of tracking, and not types of users. The motivations for self-tracking can fluctuate and overlap between these strategies.

Directive And Documentary Tracking

Directive tracking is very much targeted by goals, set either by tracking software, by training programs or by the user. Behavioural change is of course associated with this, but the study emphasises the significance of "*interweaving technologies*", stating that behaviour change is not just a possible outcome of a single technology, rather technologies that exist together in an information ecology. The challenges which result from this – specifically dispersed tracking data – will be explored over the following section.

Conversely to directive tracking, users of the documentary tracking style are not concerned with changing their behaviour. Instead they record smaller passages of their life such as memorable experiences and routine activities. Rooksby et al. discover that while some people transform observations from the documentary style into specific goals, many see it as a "*necessary*" means of occasional life observation and that documentary tracking is "*not usually a long-term endeavour*" (Rooksby et al., 2014, p.1168).

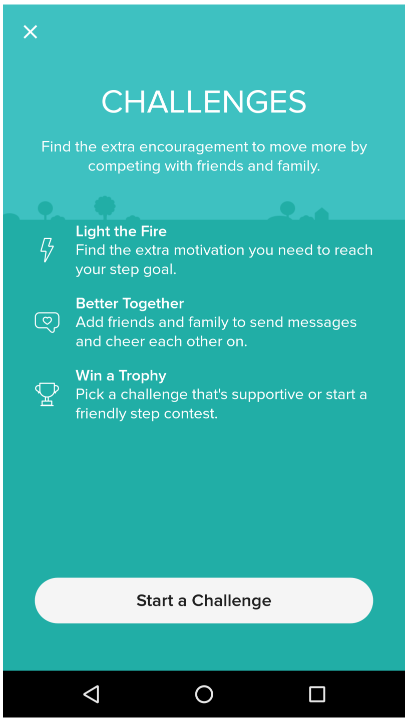


Figure 2.1: Fitbit App – Competition For Fitness Motivation

Rewards, Fetishised And Diagnostic Tracking

Related to this style is collecting rewards. Recently, gamification (Deterding et al., 2011) has emerged across new contexts, with positive implications spanning from motivation to engagement suggested from current research (Hamari et al., 2014). For several participants in their study, Rooksby et al. found a close connection between this achievement-based approach and the documentary tracking style. Indeed, the motivational affordances which gamification provides is now core functionality in many fitness related applications, including the previously discussed Fibit (see Figure 2.1) and Strava[8].

The next style, fetishised tracking, describes a style relatively synonymous with people who are "*tech savvy*", or simply early adopters. This style may have benefited from being developed alongside "*extreme users*" identified by Choe et al. (2014), enabling a broader definition of quantified-selfers under the general population.

Finally, diagnostic tracking is where the user searches for relationships between two separately tracked things, usually representative as a form of *self-experimentation*. For instance, this could entail tracking diet pattern to see an effect, if any, on sleep. Identifying meaningful correlations in personal data is a current area of research for the Human Computer

[8]https://www.strava.com/

Interaction field.

This is the focus of research by Epstein et al. (2014, p.667), who define a *cut* of data as "*a subset of the collected data with some shared feature*". Their study determines 13 cuts of location and activity data for user reflection. While this study largely focuses on physical activity, it is significant in that it is one of the larger studies ($n = 113$) that includes research around casual population motivations. Given that research into motivations so far has been limited to more extreme demographics (Choe et al., 2014) and ones that do not scale to the general population (Rooksby et al., 2014), what understanding of motivation can be gathered from this research to improve insights for casual users?

In order to develop relevant, meaningful insights for casual users, Epstein et al. (2014) completed a formative survey in which participants were questioned around their goals for self-tracking. The findings can be distributed into the tracking styles defined by Rooksby et al. (2014). The most common responses 'Maintain/Increase Activity' (41 respondents) and 'Maintain/Lose Weight' (35 respondents) fall under *directive tracking*, whereas 'Find Patterns' (7 respondents) is clearly an example of *diagnosis tracking*. It follows that the other most common responses can be categorised accordingly: 'Awareness of Activity Levels', 'Increase Motivation', 'Be Held Accountable', 'Have A Record Of Activity' and 'Competition'. The survey additionally established that respondents had an average of 1.6 goals. Principally this means that it is completely attainable for users to have a one-to-many relationship with tracking styles, taking into consideration multiple goals categorised under varying styles. Ultimately with a wide spectrum of different motivations, users will have independent goals for different types of collected data.

2.1.3 Broader Challenges Of Personal Data Collection

Businesses benefit from self-tracked data as well. Gartner (2013) predicts that consumer data collected from wearable devices will account for 5% of sales from the top global 1000 companies by 2020. The notion of the personal data economy has seen much attention in recent years and an extensive amount has been written across media outlets around its implications (Economist, 2014; Cassidy, 2016; Moody, 2016). Google's largest revenue generator – targeted behavioural advertising – is built around the wealth of information accumulated on their users. Their Android Pay API [9] allows businesses to send geofenced notifications to targeted users who are within a certain physical distance of a business storefront. It isn't unfeasible to see this extending into the wearable technology space, exploiting both the smartphone GPS sensor and the *"glanceable"* nature of smartwatch notifications to provide quick, contextual discounts based on the personal location of a user (Lyons, 2016). Similarly, Facebook allows advertisers to create targeted advert campaigns by picking from 98 personal data points of over 1 billion unique users (Dewey, 2016). With such colossal data sets centered around collected personal data, there are 3 key challenges to consider from both business and consumer perspectives.

[9] https://developers.google.com/save-to-android-pay/

"*The accumulation of personal data has an incremental adverse effect on privacy*" – Tene and Polonetsky (2012, p.251). Self-tracking injects increasing amounts of personal data into big data systems, with specific concerns around technology enabling the recall of historical data (Nunan and Di Domenico, 2013) and anonymisation (Mittelstadt and Floridi, 2016). The first challenge is associated with the privacy of personal data, which is often deeply physiological. A discussion on this continues in Appendix B, as we did not deem this challenge to fall under the scope of this project. Over the following sections we will therefore explore the two remaining challenges: device abandonment, and data dispersion.

Device Abandonment

Data sharing is cited as one of the prominent reasons for device abandonment – the second key challenge for personal data collection. Research completed by Epstein et al. (2016) questions the link between data sharing concerns and abandonment of self-tracking. Concerns were focused around location data – specifically oversharing to social media and location targeted advertising – but did not pick up on newer physiological data such as heart rate. This study extends an earlier exploration of the barriers leading to smart device abandonment completed by Lazar et al. (2015). Lazar et al. found three categories of barriers which led to device abandonment: "*devices not fitting with participants conceptions of themselves*", "*collected data not being useful*" and "*devices requiring too much work and maintenance*" (Lazar et al., 2015, p.638-640). One of the main reasons why the collected data was not useful for study participants was because of its unprocessed form. Participants did not know how to analyse and interpret the data in order to form actionable next steps.

Supporting users to better reflect on their personal data is an open area of research across the personal informatics field (Gulotta et al., 2016; Li et al., 2011). Li et al. (2011) introduce the idea of *phases of reflection*, in which people transition between *maintenance* and *discovery* phases. During the *maintenance* phase people use their collected data to maintain their behaviour relative to a goal. The *discovery* phase is where people identify data correlations to see the effect on behaviour, ultimately to identify program-level goals. Li et al. (2011, p.8) establish that the discovery phase can be analytically taxing, but that "*user involvement is critical*" for engagement with data. It follows that self-tracking tools should assist users in identifying goals.

Lee et al. (2015) investigated using a reflective goal-setting strategy to personalise a program plan. This encouraged users to deeply think about what goals they are setting, uniquely tailoring what achieving certain goals meant to them. The outcome was that this personalisation process drove users to increase their daily step count, albeit over the short-term length of the study. Gulotta et al. (2016) suggests that starting the reflection process as early as possible can play a part in acquiring longer-term benefits. Several strategies are proposed to increase engagement and create actionable next steps — using big data based upon other users to propose more realistic, achievable goals, and providing the user with more tailored, achievement-based feedback if they begin to stray from their goals.

A further reason attributed to abandonment is the physical cost of data collection and maintenance (Epstein et al., 2016; Lazar et al., 2015). While automation in smart devices can help to reduce the manual effort associated with spreadsheets, both studies noted the hassle which participants felt in maintaining devices. It is clear from Lazar et al. (2015) that there is an inherent link between the value gained from self-tracking versus the time expensed doing so. Participants simply stopped tracking due to hassle, laziness or loss of interest (Epstein et al., 2016). An additional barrier specified by (Lazar et al., 2015, p.638) is that self-trackers did not find the data collected useful — "*they were not interested in the level of information the data gave them*". Therefore, research has looked towards including additional self-quantified data in order to find "*more appealing inferences*"(Haddadi and Brown, 2014, p.1). This brings us to the third challenge.

Data Dispersion

As previously discussed, Rooksby et al. consider that behavioural change may be the outcome of "*interweaving technology*". With diagnosis tracking being a common goal for self-trackers, how can users best explore data correlations in dispersed, self-tracked data sets? The arrival of aggregated data dashboards aim to tackle this question (e.g Exist.io[10], Zenobase[11], Fluxtream[12], TicTrac[13]). With the increasing ubiquity of smart devices, sensors are accumulating more data than ever before. However, this data is scattered across all types of distinct devices, open APIs and closed platforms, with no single tool supporting the tracking and exploration of all data. In the case of extreme quantified-selfers, Choe et al. (2014) identify that people who have the technical ability build their own tools for analysis and self-experimentation to remedy this problem.

Existing research also points to users simplifying tracking strategies due to being unable to extract the significance of high-dimensional data (Choe et al., 2014). Clearly there is a space for aggregated platforms like Exist.io to appeal to broad audiences, offering a comprehensive, system-driven method of exploring self-tracked data. The benefits of these platforms are numerous. Self-tracking services such as Fitbit and Last.fm are integrated into the platform automatically. Exist.io then analyses data across distinct services, automatically investigating relationships between the users chosen connected services. Finally, visualisations are continuously computed to aid non-experts in the reflection process.

Jones and Kelly (2016) highlight the sensemaking challenges associated with aggregated self-tracking platforms such as Exist.io. In particular, the cognitive effort required to examine vast quantities of data alongside the correlational nature of the data itself. One participant in their study experienced "*analysis paralysis*" (Jones and Kelly, 2016, p.3)— consistent with open challenges identified by Choe et al. around the steep learning curve users may experience when first faced with analysing data visualisations. A further sense-

[10]https://www.exist.io

[11]https://www.zenobase.com

[12]https://www.fluxtream.org

[13]https://www.tictrac.com

making challenge focuses on the difficulty of inferring holistic insights from disjointed outputs (Jones and Kelly, 2016). In this context, the aggregation platform needed to provide explicit, user-driven functionality to chain multiple dispersed datasets. Much like how Choe et al. envision achieving synergy between user-driven and system-driven data collection, this is also true of data reflection.

A mixed-initiative approach (Horvitz, 1999) should be explored for personal informatic systems, supporting the discovery of powerful data insights and maximised user engagement. This approach, amongst others, will be considered over the remainder of this chapter as methods of reflecting upon collected personal data. A discussion of visualisation techniques firstly occurs in the following section, and finally we explore the potential of virtual reality as a medium for immersive personal data exploration.

2.2 Visualisation

2.2.1 A Mixed-Initiative Approach To Exploring Data

Allen et al. (1999) discuss the duality of AI and HCI and how strategies between these fields leads to mixed-initiative interaction. Adjusting to the current context, either human agents or system agents will contribute towards a system at a suitable time. Data exploration may leverage a mixed-initiative approach to support increasing the user's understanding of the data. This deep means of collaboration has been debated previously (Shneiderman and Maes, 1997), with a flexible interaction strategy between agents providing automation, and agents directly interfacing with the system, cited as the goal for a combined interaction approach (Allen et al., 1999). Visualisation techniques are then layered on top of this approach to minimise the cognitive load of users (Pu and Lalanne, 2002). Clearly the complexity and diversity of visualisations will widely range between systems and their intended audiences.

Nevertheless, both human and intelligent agents have respective limitations when it comes to supporting visualisations. Modern computers are generally scalable, and can handle increasingly large simulations and data transformations. Humans may not be as computationally powerful in this area, but are experts at interpreting the results using natural "*highly developed pattern-recognition skills*" (Van Dam et al., 2000, p.27). The measure of an effective visualisation is when humans are able to identify unpredicted trends and data characteristics (Saraiya et al., 2005). Measuring this form of effectiveness will therefore be a key evaluation criterion for this project.

Epstein et al. (2014) explore the value of discovering unexpected data, potentially patterned data, in a contextually aware environment. Indeed, the timing and location of data visualisation and the insights which humans can gain from them is an important engagement factor. Intille (2004) investigates the opportunities of providing *just in time* information at points of decision, finding that behaviour change can be successfully achieved in non-technological fields. Therefore the aim of visualisation is to communicate information clearly such that insights can be inferred from data features and patterns, and also to enable viewers to ask questions of visualisations to identify new hypotheses (Saraiya et al., 2005).

2.2.2 Presenting Data Using Dashboards

A traditional dashboard paradigm is often employed as a starting point for data exploration. Few (2013) offers the following definition for a dashboard:

"*A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at glance*"

The glanceability of dashboards cannot be overstated; conveying a high-level overview of data reduces the burden of information overload. Data summaries are concise and are primarily used to open the door to lower-level details. This extends itself to working human memory which supports a finite number of *visual chunks* — dashboard design should enable efficient understanding of information in *visual gulps* (Few, 2013, p.79). Ultimately, dashboards transform data repositories into consumable information (Hovis, 2002), equipping users with a tool to deduce visual patterns and make effective decisions (Brath and Peters, 2004).

Figures 2.2 and 2.3 show the companion app dashboard for a Fitbit tracker. The dashboard contains both text and graphics, but is primarily graphical. It acts as a visual overview of activity, with a particular emphasis on visual — approximately 70% of all sense receptors in the human body are situated in the eyes (Few, 2013). Consistent with Few's definition, the means of achieving objectives are clearly stated: e.g the coloured, circular progress bars for physical activity in Figure 2.2 and "*500ml left*" for the water target in Figure 2.3. Clicking on activities opens an exploratory environment where users can analyse the chosen activity in greater detail. For example, clicking on the heart symbol in Figure 2.3 opens a new screen dedicated to exploring the user's heart rate statistics in Figure 2.4. Inconsistent with Few's definition, the dashboard is not consolidated to a single screen. However, it is worth noting that Few talks of the compromises to be made in sacrificing "*effective design to accommodate diversity in screen types*"(Few, 2013, p.70).

Figure 2.2: Fitbit Dashboard: Above the fold

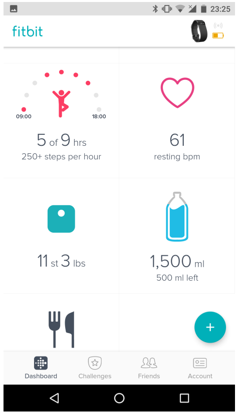


Figure 2.3: Fitbit Dashboard: Below the fold

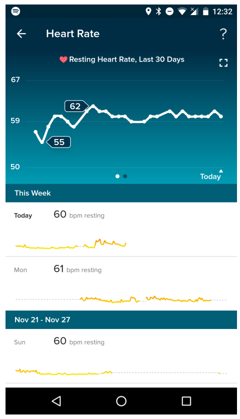


Figure 2.4: Fitbit: Lower-level analysis

Natural Language For Visualisation

Bentley et al. (2013) consider a modern skill gap across developed countries such as Germany and America in which general audiences have difficulty grasping simple bar and pie charts. Consequently Bentley et al. explore an alternative method of communicating and presenting statistical data to users; through natural language. Statistically significant results were displayed to study participants with the neutral form "*On days when you X, you Y*" (Bentley et al., 2013, 30:4). Similarly, short verbal data summaries have also been evaluated by Epstein et al. (2014) who presented them alongside data visualisations to their study participants. In both cases, the use of natural language was found to create behaviour change opportunities. Exist.io is a commercial example which uses the combination of natural language statements and visual graphs to present correlations in aggregated self-tracked data (see Figure 2.5).

There are challenges associated with using natural language to explain data correlations. Li et al. (2010, p.6) establishes that most personal informatic systems "*do not have specific suggestions on what do next*" when people reflect on their data. Indeed, this was extended by (Epstein et al., 2014) who found that people expected actionable steps when reflecting with the use of natural language. In particular, participants who had difficulty interpreting graphical visualisations expected natural language insights to serve as both "*an explanation and identification of actionable changes*" (Epstein et al., 2014, p.9).

2.2.3 Filtering For Meaningful Insights

Moreover, discerning between correlation and causation for a relationship between two variables can be difficult task for those reflecting on their self-tracked data. This ties into the reality that not all outputs of personal informatics systems will be insightful, often due to the virtue of them being obvious. Jones (2015) considers the task of filtering vast

Figure 2.5: Exist Correlational Data

quantities of self-tracked data to select the most interesting insights. Manual analysis by users is unfeasible, given that additional correlations will emerge as dimensions are added to the system (Jones, 2015). Hence, Jones and Kelly (2016) examine introducing automation into the process for filtering for interesting correlations. However, a fully-automated approach based on generalised participant-based criteria for *interestingness* did not yield significant value for participants. Therefore, Jones and Kelly suggest a different approach such as semi-automated filtering, where user and system collaborate together to capture the user's measure of interestingness on a deeper, individual level. Clearly, there are many open questions around determining the most interesting correlations to present to users, and furthermore, how systems can best describe data correlations and provide actionable next steps.

Casual Infovis Distinctions

Pousman et al. (2007) define *Casual Infovis* to cover a broad practice of visualisation for a wider audience. This term complements traditional information visualisation and encompasses many other types including ambient, social and artistic visualisation. Pousman et al. describes the usage patterns of traditional infovis systems as *episodic* — domain experts will deeply explore datasets over a number of hours. Alternatively, Casual Infovis users will dip in and out with "*fleeting moments of inspection*" (Pousman et al., 2007, p.1149). Since much of Casual Infovis illustrates personal data, Pousman et al. (2007, p.1151) note that visualisations are increasingly "*meaningful to users*", comparative to the data in traditional systems purely being "*efficient and effective*".

Stusak et al. (2014) investigate data physicalization, a subset of Casual Infovis, where in their study running activity was represented through a physical form. Their study discovered that personalised physical sculptures encouraged participants to interpret and reflect on their activities — fostering engagement by motivating users to self-experiment to change the sculpture's shape. Evidently the objectives of Casual Infovis systems are different to those of traditional visualisation systems, and accordingly the insights which we gain from these systems are also different.

2.2.4 Towards Visualisation Techniques

Although there is much overlap between the two disciplines, there is an important distinction to be made between scientific visualisation and information visualisation. According to Card et al. (1999, p.6), scientific visualisation is the "*use of interactive visual representations of scientific data, typically physically based, to amplify cognition*". Information visualisation also shares the objective of amplifying cognition, but through an interactive visual representation of "*abstract, non-physical data*" (Card et al., 1999, p.6). Munzner (2008) extends this definition by stating that spatial representation is *chosen* for information visualisation, whereas it is *given* for scientific visualisation. These definitions reflect that information visualisation makes design choices on how best to abstract data for a chosen task. Nevertheless, for the purposes of this project and with this distinction considered, after this section both of these terms will be unified under the umbrella term of *data visualisation*.

Immersive Scientific Visualisation

Scientists have looked towards visualisation to "*enhance comprehension and deepen insight*" since requirements to analyse large, complex data sets emerged in the 1970s (Haber and McNabb, 1990, p.74). Visualisation plays a crucial role in interfacing between the human visual system and the vast computational power of computers (Ware, 2004, p.2). These cognitive systems support the understanding of underlying datasets in scientific disciplines ranging from molecular biology to astronomy. Ware (2004) highlights that a view of both small-scale and large-scale features of the data can be observed with visualisation, allowing the perception of significant patterns at both a local and global level. Techniques for mapping data to 2D visualisations are now widespread, with popular software for generating diagrams and graphs including matplotlib[14], Tableau[15], Google Charts[16]and D3.js[17].

High-dimensionality structures place a considerable demand on real-time visualisation of large datasets, particularly with the larger datasets scientific visualisation typically contends with (Bryson, 1996). Indeed, notably as the dimensionality of data increases, so does the complexity of user interfaces which support the exploration of this data (Steed et al., 2016). Scientists have therefore evaluated solutions within Immersive Analytics[18] such as the CAVE2 virtual environment[19] — a 320 degree circle of 72 LCD panels with multiple

[14]http://matplotlib.org/
[15]http://www.tableau.com/
[16]https://developers.google.com/chart/
[17]https://d3js.org/
[18]http://immersiveanalytics.net/
[19]https://www.mechdyne.com/hardware.aspx?name=CAVE2

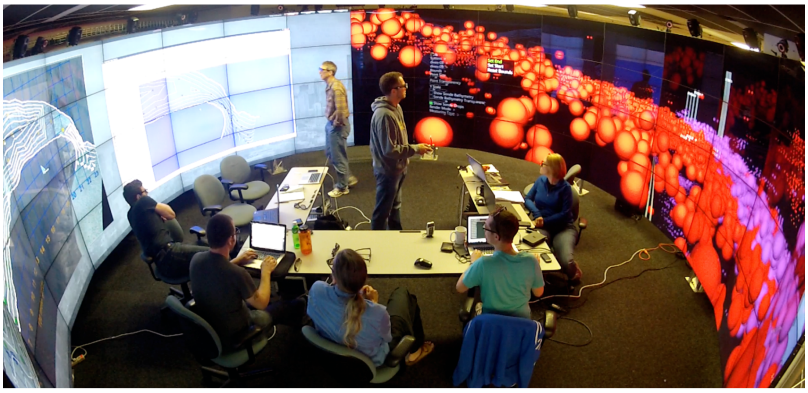


Figure 2.6: Visualisation of large datasets on a CAVE2 system (Marai et al., 2016)

tracking cameras – to provide new opportunities for sensemaking.

CAVE is an example of an adaptable, immersive visualisation system for large scientific datasets. CAVE-like systems offer great promise towards collaborative data exploration, but there are still challenges ahead for this strand of visualisation (Marai et al., 2016). Evidently manual data integration and programming the visualisation process is suitable for experts only and the system is only typically concerned with scientific data, which is not within the scope of this project. However, the research around interactivity within this form of cognitive system forms part of the wider context surrounding interaction and dimensionality.

Keim (2001, p.40) states that visual representations of data provide a "*higher degree of confidence in the findings*" of data exploration than standard textual representations do. It follows that multiple visualisation techniques will be required to achieve this for different types and dimensionalities of data. For instance, visualisation of a simple x-y graph plot versus grouped data with values on one dimension represented through colour. There is not a universal solution and consequently evaluation of distinct visualisation techniques must take place on an individual basis for the task at hand (Keim, 2001). As a result, much of the research around particular visualisation techniques is task and domain specific.

A Taxonomy For Visualisation

Building upon previous taxonomies devised by Chi (2000) and (Keim, 2001), Chengzhi et al. (2003) proposes a visualisation taxonomy which takes the viewpoints of both users and developers into account. Two separate frameworks were created, targeting the correct application of visualisation techniques fron both user and developer perspectives. For instance, the user-oriented framework considers *Text, 2D, 3D and multi-dimensional* data types amongst several others. Specific examples of visualisation techniques include a Perspective Wall (*Text*), a scatter plot (*2D*) and Grand Tour (*multi-dimensional*).

Comparatively, the developer-oriented framework views the relationship between representation modes such as *pixel-oriented* and *geometric projection*, and states corresponding interaction levels ranging from manual to automated analysis for the visualisation operator. The purpose of this taxonomy is to support the correct application of visualisation techniques to appropriate domains, and indeed to highlight research gaps and challenges across the frameworks.

A thorough search of personal visualisation literature and applications yielded many examples of 2D visualisation techniques, but no examples of 3D visualisations for representing personal data. Within the personal informatics field, the lack of 3D visualisations is one such example of a research gap.

Choe et al. (2015) analysed the frequency of specific visualisations during presentations at a Quantified-Selfers meet-up. The most frequently used visualisation techniques amongst participants included a line chart, bar chart, and scatter plot. It is notable that even with a more extreme demographic of self-trackers, traditionally these visualisations are used in 2D environments, and that no examples of 3D visualisations were recorded by Choe et al.. Indeed, searches of current literature all showed users interacting with these forms of personal data visualisations in 2D, with some alongside natural language (Bentley et al., 2013; Epstein et al., 2014; Choe et al., 2014; Jones, 2015). A 3D design space is "*self-evidently richer*" than its 2D equivalent, but this does not automatically result in 3D providing more effective access to information (Ware, 2004, p.294). Given that we have established that many visualisation techniques are domain specific, further research is needed within the context of personal data to measure the effectiveness of moving from 2D to 3D visualisations for the general population.

Dealing With Dimensionality

During this project, one of our objectives is to explore ways in which distinct, dispersed datasets can be combined and visualised. The integration of these separate datasets will result in an increased dimensionality. However, traditionally humans have trouble perceiving high-dimensional data – a conceptual barrier which causes difficulties with "*proper intuition of the properties of high dimensional space*" (Jimenez and Landgrebe, 1998, p.2).

One solution to increases in dimensionality is to adapt low-dimensional visualisations to represent further dimensions within the same dimensional space. For instance, using attributes of existing data – such as colour and shape – to depict an additional dimension. Indeed, a different visualisation technique can be used altogether, such as a radar chart or parallel coordinates as seen in Figure 2.7, which allow visualisation of high-dimensional data.

Alternatively, the data itself can be adjusted in a process known as dimensionality reduction. One part of this process is feature selection, the exercise of acquiring a subset of relevant, informative features in the data. The second part is feature extraction, the process of projecting highly-dimensional data into a lower-dimensional space. Conventional

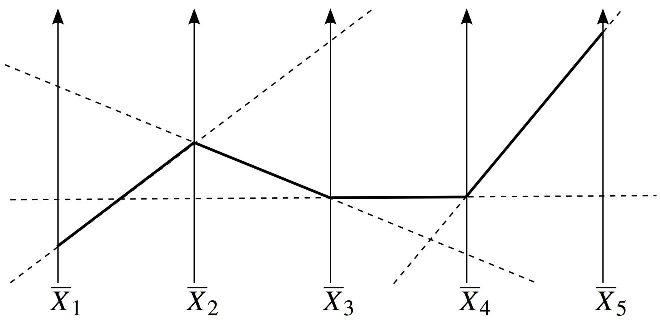


Figure 2.7: Parallel Coordinates with N=5 dimensions (Heinrich and Weiskopf, 2013, p.98)

methods such as Principal Component Analysis are traditionally utilised for extracting new features. Both of these methods are therefore beneficial for improving computational properties and for reflecting upon the most important characteristics of high-dimensional datasets efficiently (Cunningham and Ghahramani, 2015).

2.2.5 Parallel Planes and Be The Data for Visualisation Insight

We have identified two different visualisation techniques which we will evaluate throughout our study. The first technique – Parallel Planes – enables high-dimensional datasets to be represented in a single visualisation. Be The Data is the second technique and lets users navigate around a three-dimensional scatter plot, and take on the perspective of data points. The aim of our study is to see whether both of these techniques support users to understand and reflect on personal datasets in an immersive environment.

Parallel Planes

Brunhart-Lupo et al. (2016) extends the aforementioned parallel coordinates visualisation technique by mapping multivariate data into a three-dimensional space. By mapping onto a series of parallel rectangles, rather than a series of parallel planes, and exploring the results within an immersive environment, the technique helps to alleviate the over-plotting problem associated with standard parallel coordinates and supports data analysts in locating correlations in the data (Brunhart-Lupo et al., 2016). An example visualisation can be seen in Figure 2.8.

In the context of this project, Brunhart-Lupo et al.'s study raises two important questions:

- Can immersive environments be used as a medium for effective personal data exploration? Brunhart-Lupo et al. cite existing literature describing the benefits of immersive environments for analysing data, but much of the research focuses on sci-

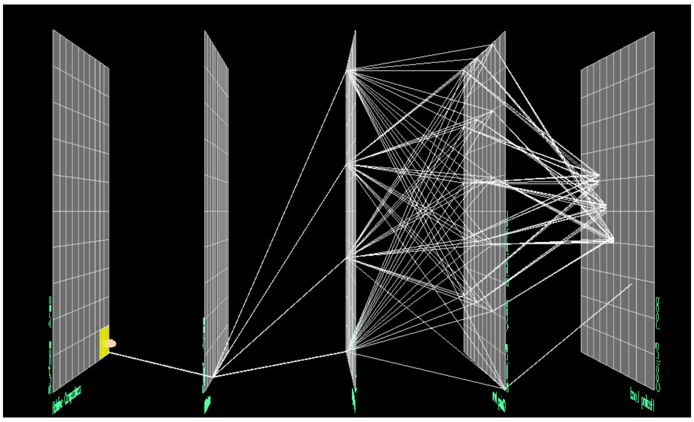


Figure 2.8: Immersive Parallel Planes Visualisation (Brunhart-Lupo et al., 2016)

entific visualization (Henry and Polys, 2010) or more broadly (Arms et al., 1999), without exclusively looking at personal data.

- Are potentially complex high-dimensional visualisations such as Parallel Planes a suitable visualisation technique for non-expert users? The users in Brunhart-Lupo et al. (2016) were simulation analysts, so it is not yet clear whether this technique is feasible for the wider population.

A further area explored by Brunhart-Lupo et al. is brushing. Brushing is a mechanism for "*interactively selecting subsets of the data*", which then enables further actions such as highlighting or deletion (Martin and Ward, 1995, p.271). Explicitly focusing on specific observations, particularly in a parallel coordinates scenario, allows the user to detect correlations across multiple dimensions (Hauser et al., 2002). The significance of brushing in this context was the ability for analysts to generate new simulations on selected data regions. This links back to Saraiya et al. (2005, p.1), who assert that visualisation not only enables data exploration, but "*to also find questions that identify new hypotheses*".

Be The Data

Techniques for visualising global and local perspectives of data have previously been expressed as one of the top scientific visualisation research problems (Johnson, 2004). *Be The Data* by Chen et al. (2016) discusses an educational approach to teaching students about data analytics tasks such as dimension reduction. Rather than purely viewing the data from a global perspective, students embodied and became virtual data points. Students collaborated in a physical environment and motion-tracking technology was used to track students across the room. As they moved, the corresponding data points moved with them in the original projection of data. The results from the study were positive — users were engaged, "*exploiting embodiment*" to further analyse the data in new ways (Chen et al., 2016, p.5). Therefore, extending the *Be The Data* visualisation approach will be a focus for this project. More specifically, the notion of letting users take on the perspective of a data point. However, rather than this taking place in a collaborative physical space, the notion will be applied on an individual basis in an immersive, virtual environment.

Choe et al. (2015) categorise eight data-driven insight types to inform the design of data exploration tools: *Detail, Self-Reflection, Trend, Comparison, Correlation, Data Summary, Distribution* and *Outlier*. Previously, literature has defined insights across non-personal data domains — however, Choe et al. directly examine insights within the context of personal data elicited from Quantified Selfers. Although these insight categorisations may not extend to representing the general public, Choe et al. argue that this community is at the forefront of understanding personal data visualisation and can nonetheless provide a valuable understanding of insights.

Using this model, the Parallel Planes technique is an example of a *data summary* insight type, and the Be The Data technique represents a *detail* visualisation insight. The principal purpose of visualisation is insight (Saraiya et al., 2005; North, 2006), and therefore understanding the different forms of insight types is beneficial towards creating more powerful, insight-generating visualisation platforms. It has been argued that immersive visualisation multiplies "*the effectiveness of desktop visualization*" (Donalek et al., 2014, p.610), and that emerging immersive technologies harness the "*remarkable pattern recognition systems*" which humans possess to gain more intuitive data understandings (Donalek et al., 2014, p.609). Indeed, immersive environments contain many desirable qualities – global context, easier navigation and adventurous interaction amongst others (Van Dam et al., 2000) – which can augment the capacities of humans to interpret complex datasets. The following section therefore investigates the potential of using a fast-growing immersive platform, Virtual Reality, as a means for perceiving insight through personal data visualisation.

2.3 Virtual Reality

This section gives an overview of the virtual reality field. We begin by defining the term and contextualising its use across scientific and healthcare domains. The distinction between immersion and presence is stated, and the immersive characteristics of the platform are detailed. Finally, we discuss the opportunities of using a smartphone for both collection of personal data, and reflection of personal data in VR.

2.3.1 Backdrop of VR

Virtual reality (VR) is undergoing a resurgence. VR first rose to prominence in the 1990s with Sega's VR project and the Virtuality gaming machines – the primary breakthrough products. However, VR never took off then for consumers, with the effects of nausea, weak hardware and the lack of global standards cited as reasons for its early decline (Horowitz, 2004; Arthur, 2015). Two decades later, ubiquitous smartphones have not only revolutionised many aspects of daily life, but also have driven a demand for hardware miniaturisation. With the advent of tiny, immensely powerful chips, there has now been a shift back towards VR – largely propelled by immersive gaming experiences which first brought VR to fame in the 1990s. Importantly, in late 2016 an open consortium was established, consisting of large technology companies such as Google, Intel, Samsung and Nvidia. The Khronos VR standard[20] will enable portability between different VR systems, and will include open APIs for a wide range of VR functions. Ultimately the purpose of this standard is to enable further growth within the VR market – of which analysts have predicted will possess an addressable market size in 2025 larger than today's television market (Bellini et al., 2016).

Firstly we must define what VR is. When VR first emerged in the 1990s, many early definitions naturally conflicted and were tied to specific hardware, rather than addressing the forms of human experience which VR produces (Steuer, 1992). In their seminal paper, Cruz-Neira et al. (1993) adopt the definition that VR is an experience in which a person is "*surrounded by a three-dimensional computer-generated representation, and is able to move around in the virtual world and see it from different angles, to reach into it, grab it, and reshape it*" (Cruz-Neira et al., 1993, p.1). While this definition is hardware agnostic, Steuer (1992) also argues that VR must be thought of in terms of perceptual factors, notably "*mindful attentional, perceptual, and other mental processes*"(Steuer, 1992, p.6). A succinct summarisation of all of these processes is *presence*, which shall be discussed over the following sections. Therefore, for this project we will define VR in line with Bryson's definition, who describes VR as: "*the use of computers and human-computer interfaces to create the effect of a three-dimensional world containing interactive objects with a strong sense of three-dimensional presence*" (Bryson, 1996, p.62).

2.3.2 Disciplines And Applications For VR

Van Dam et al. (2000) reported on an accelerating data size crisis in the scientific community where scientists ability to interpret and visualise data is far outweighed by their ability to produce and collect data. With management prioritising investment in data computation over data visualisation technology, VR is considered to be a galvanizing technique to strengthen the bond between visualisation and human understanding of large datasets (Van Dam et al., 2000). Through a multidisciplinary approach, scientists have leveraged the power of existing developments across gaming industries in the absence of substantial visualisation investment (Bryson, 1996), and the benefits of VR are ample. Virtual reality enables global context through "*much more use of peripheral context*" (Van Dam et al., 2000, p.32), allowing the human eyes to consume a greater amount of information within the data exploration environment. Indeed, VR is an interface which offers "*significantly enhanced three-dimensional perception*" (Bryson, 1996, p.64), distinctly lending itself to the scientific demands of three and higher-dimensional data exploration. Artificial intelligence will support scientific understanding in the future, but clearly the fundamental appeal of

[20] https://www.khronos.org/news/press/khronos-announces-vr-standards-initiative

present VR technology for scientists is for gaining "*rapid insight*" (Van Dam et al., 2000, p.35) of expanding, complex scientific datasets.

VR also has important use cases across healthcare (Bellini et al., 2016; Mantovani et al., 2003), largely relating to education and training. Interaction is the central element to VR experiences which encourages active participation, rather than passivity, during learning (Mantovani et al., 2003). This empowers students to understand, shape and learn from the virtual world in their own context, which culminates in "*more meaningful and effective learning*" (Mantovani et al., 2003, p.390). Consequently, there are a multitude of medical VR applications, such as the orthopedic surgery simulator (Tsai et al., 2001), which look to address task-specific clinical skills. Seymour et al. (2002) showed that it is possible to train surgeons to specified objectives through VR and successfully transfer these skills to patients in real operating room environments. Primarily these forms of VR applications come closer to real-world experiences than conventional 3D desktop-based applications do (Van Dam et al., 2000), making them compelling educational tools for fostering engagement with users.

Much like education and training, VR can be used for cost-effective solutions beyond the healthcare field. There are applications in simulation and verification of manufacturing processes (Ong and Nee, 2013), as well as the architectural designs of buildings and landscapes (Portman et al., 2015) – two examples of potentially inaccessible realities attainable with VR. Using a "*lifetime of experience*" of making spatial decisions, humans can navigate more effectively through VR in comparison to traditional desktop based environments (Van Dam et al., 2000, p.32). Head-tracking – one of the fundamental characteristics of VR – allows for objects to remain in a fixed position while the user explores the environment.

This form of organic navigation can be beneficial in scenarios such as architectural walkthroughs, where designers wish to experience their plans before they are physically built. In contrast, input devices such as a mouse, keyboard or joystick are used with 3D desktop environments to move the environment around a fixed user view. Consequently, navigation is more natural in VR and furthermore supports improved search task performance (Pausch et al., 1997). Pausch et al. (1997) shows that VR supports users to remember previously viewed locations – an important instinctual memory recall technique for pilots training in cockpit simulators, for instance. VR therefore has interdisciplinary characteristics which are widely applicable over a wide-range of use cases.

2.3.3 What Does VR Offer? What Makes It Engaging?

There is one characteristic of VR which is a focal target throughout all of the disciplines. This is *presence*. Steuer (1992, p.6) defines presence as "*the sense of being in an environment*" – in effect, the perceived reality of objects in the VR experience and the users' presence with these objects (Van Dam et al., 2000). In this context, presence is principally the users' relative distinction between them being in scenes in virtual and real world environments (Schuemie and Van der Mast, 1999). The notion of presence will be used across this project as a measurable condition of a users' subjective experience in VR.

Establishing presence is an area of interest for researchers and applications which harness VR. One of the earliest examples by Hodges et al. (1995) explored invoking presence as an exposure therapy technique for Acrophobia[21]. Much research on exposure therapy and presence has followed since, particularly for anxiety disorders such as Arachnophobia[22] and Aviophobia[23] (Powers and Emmelkamp, 2008). Riva et al. (2007) demonstrates the bidirectional relation between presence in VR and an individual's emotion, although the study was limited to two emotional states of anxiety and relaxation. Indeed, the full extent of the relationship between presence and emotion is somewhat undefined and represents a research gap across the psychology field (Diemer et al., 2015). Nevertheless, within clinical psychology, Hodges et al. (1994, p.10) asserts that presence was the "*defining factor in the success*" of their exposure therapy application.

To understand presence, first we must consider the components which it is composed of. Sheridan (1992) proposed three orthogonal principles that construct presence. These are:

- Extent of sensory information
- Control of relation of sensors to environment
- Ability to modify physical environment

Each principle can be varied independently, as reflected in the three labeled axes in Figure 2.9. *Perfect presence* is an unspecified function of the combination of axes, with *lines of constant information flow* indicating that the *extent of sensory information* consumes more information than the two related but independent control components (Sheridan, 1992). The independence of the determinants lead Sheridan to clarify them to be task-dependent – both in terms of the task difficulty, and the degree of task automation. The level of automation between a manual or automatic approach (perhaps resulting in a mixed-initiative interaction) will affect the user's perceived presence and/or performance ability.

Distinguishing Immersion And Presence

A distinction between immersion and presence must also be made. Slater and Wilbur (1997, p.606) state that immersion is an "*objective and quantifiable description*" of what systems can provide, whereas presence is a "*state of consciousness*" of the user. Therefore, immersion describes the characteristics of a technology, and the *sense of presence* it builds for a user. Slater and Wilbur (1997) further breaks down immersion into four areas: *Inclusive*, *Extensive*, *Surrounding* and *Vivid*. This categorisation of immersion goes hand in hand with Sheridan's model of presence, but notably there is not a direct relationship between them. Cognitive processes which sit between immersion and developing presence can become a hindering factor (Schubert et al., 2001). Furthermore, there is a

[21] The fear or phobia of heights
[22] The fear or phobia of spiders
[23] The fear or phobia of flying

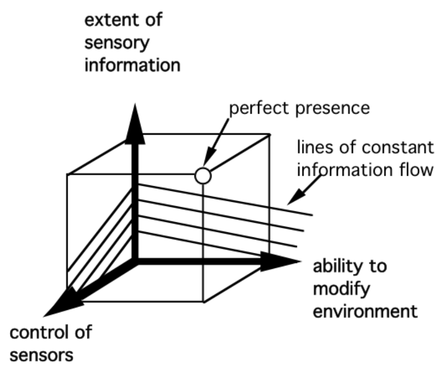


Figure 2.9: Sheridan's three principal determinants of presence (Sheridan, 1992)

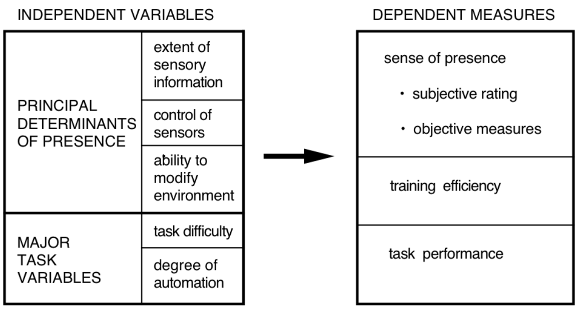


Figure 2.10: Sheridan's experimental determination of presence, learning efficiency and performance (Sheridan, 1992)

strong dependency on a user's ability to operate a user interface which Slater and Wilbur (1997) describe as crucial with regards to task performance. As a result, presence and task performance are not necessarily analogous. Sheridan's dependent measures in Figure 2.10 accordingly consider presence and task performance independently.

Inclusive Immersion

Inclusive immersion refers to the extent to which the virtual environment is occluded from the physical environment. Factors which may impact inclusive immersion include the physical properties of the VR devices themselves. Headsets can be uncomfortable to wear, particularly for prolonged periods of time, and protruding wires from the headset can negatively affect the VR experience. In the context of scientific visualisation, Van Dam et al. (2000, p.38) assert that VR will not become a normal part of an office environment until it "*literally becomes indistinguishable from that environment*". Additional challenges may arise with sound or other sensory interference from the physical world, highlighting the importance for an immersive system to maximise its influence along the *extent of sensory* *information* axis defined by Sheridan.

Slater and Wilbur (1997) define *matching* – head-tracking – separate to inclusivity, but arguably this definition is a naturally ingrained component of inclusive immersion. Head-tracking is a pervasive characteristic of modern VR systems, with numerous studies demonstrating its positive effects on presence (Slater and Wilbur, 1997). Without head-tracking, the virtual environment is not able to respond to physical head motion, breaking one of the major interface goals in VR which is to create "*natural, human-like interaction*" (Van Dam et al., 2000, p.41).

Extensive Immersion

Extensive immersion indicates the collection of sensory features which the systems supports. This covers all three axes in Sheridan's model of presence. Sound is one such example of a powerful medium for human perception – for instance, the aural rendering of system alerts, spatial cues, and the indication of objects outside of the visual display. VR can also use three-dimensional spatial sound to simulate distance and the geometry of the user's virtual environment, which has been found to increase a users' sense of presence (Slater and Wilbur, 1997).

Sound can be used alongside haptic feedback, with other forms of interaction such as simultaneous speech and hand gestures driving the system. Van Dam et al. (2000) argues that multimodal interaction is a far richer way of interacting with an environment than a mouse and keyboard. Consequently, VR can be seen as "*a natural extension of existing computing environments*" (Van Dam et al., 2000, p.27). VR exists at one extreme of the spectrum, beginning with a keyboard and a text-only display, moving to 2D graphics and adding a mouse, and then proceeding to 3D desktop graphics with joysticks, finally finishing with multi-modal interaction in VR (Van Dam et al., 2000).

Surrounding Immersion

Surrounding immersion represents the extent to which the system is "*panoramic rather than limited to a narrow field*" (Slater and Wilbur, 1997, p.605). Immersion is enhanced by a wider field of view comparative to traditional desktop displays, enabling the consumption of extra peripheral information. This is beneficial for "*situational awareness and context*", which supports navigational and spatial decision-making, ultimately augmenting presence (Van Dam et al., 2000, p.27).

Vivid Immersion

The final category of immersion is *vivid* immersion. This depicts the "*richness, information content, resolution and quality of the displays*" (Slater and Wilbur, 1997, p.605). CRTs and LCD panels were one of the limiting factors of previous generation VR technology, largely due to poor screen resolutions and physical property restrictions. Advances in screen types to OLED and AMOLED technologies have resulted in resolutions up to 2160x1200 on popular devices such as the Oculus Rift[24] and the HTC Vive[25]. At this screen resolution the limitations are the processing hardware which drives the display, as well as battery life on smartphones supporting VR experiences.

The content itself is also an important aspect of vivid immersion. According to Witmer and Singer (1998), the greater the consistency between information in a virtual environment and that learned through the real-world, the greater the presence shall be. There is also a small amount of evidence to suggest that photo realism – e.g environment illumination with reflections and dynamic shadows – is associated with gains in reported presence (Khanna et al., 2006). However, a major constraint with Khanna et al.'s study was that the VR system only delivered a low 15fps (frames per second) to keep a stable frame rate, which may have affected the findings. Modern VR platforms such as Google Daydream[26] and Playstation VR[27] specify a stable minimum of 60fps. This display requirement is also closely related to *inclusive* immersion, and ensures that immersion is not disrupted by frame rate should the user move their head in motion.

Related to frame rate is the latency of the VR experience. System latency is exceptionally important in VR systems where poor latencies can cause an interruption in immersion, and perhaps even motion sickness. Van Dam et al. (2000, p.41) establishes that latency doesn't result in sickness until hitting a "*task and user-dependent threshold*". Particularly with head-tracking (as opposed to hand/body-tracking which is relatively less demanding), latencies must be kept to an absolute minimum. Van Dam et al. cite 35ms to avoid a mismatch in visual and interaction cues. Consistent and minimal latencies enable easy transitions between cognition and perception (Van Dam et al., 2000), allowing for effective task performance in virtual environments.

Clearly presence is a multi-sensory experience, with a great deal of additional influential factors as defined through the categorisation of immersion. The purposefully unspecified function of *perfect presence* in Sheridan's model of presence is a popular research area, with a full understanding not yet forthcoming. We have defined factors known to increase both immersion and presence, as well as discerning between the two terms, and how task performance is independent of them. However, this discussion has only taken place in the context of a virtual environment.

Figure 2.11 shows the classifications of reality through to virtuality. We have considered immersion and presence at *virtual environment*, with the rest of the continuum out of scope for this project. However, there is active research across the rest of the continuum – in particular, augmented reality (AR) in which reality is virtually modified by a computer. The Microsoft Hololens[28] is a mixed reality device which uses a multitude of sensors,

[24] https://www.oculus.com/
[25] https://www.vive.com
[26] https://vr.google.com/daydream/
[27] https://www.playstation.com/en-gb/explore/playstation-vr/
[28] https://www.microsoft.com/microsoft-hololens

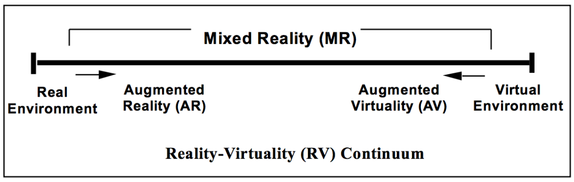


Figure 2.11: Reality Virtuality Continuum (Milgram et al., 1995)

cameras and microphones to augment the user's experience with a physical environment. As a result, the forms of presence and immersion differ in this experience, and across the continuum itself. Indeed, Baos et al. (2000) describe experiences in which people have assigned reality judgments but do not feel specifically present, and vice versa. Nevertheless, VR devices have been released in greater numbers across the consumer market relative to their AR counterparts up until this point, and as a result are the focus of this project.

2.3.4 Towards Mobile VR For Personal Data Visualisation

The gaming industry has driven the growth of head-mounted displays (HMD) such as the Oculus Rift, HTC Vive and Playstation VR. These VR devices connect to a computer via an HDMI cable which controls the visual display on the HMD. Internal sensors enable low-latency head-tracking, and a stereoscopic image is created for the users with the inclusion of lenses inside the HMD. Motion tracking can be supported through additional accessories such as external sensors positioned around the physical environment, smart gloves and wireless controllers.

These VR devices are popular with gamers due to their ability to connect with existing hardware and provide immersive, high frame-rate experiences. However, these are still early generation devices and challenges such as protruding cables and the necessity of buying additional hardware for increased immersion do exist. Furthermore, the total cost and requirements needed for high-end VR systems can be prohibitively expensive – Nvidia reported that less than 1% of computers globally in 2016 would be able to run VR technology (BBC, 2016).

Google Cardboard[29], Google Daydream and Samsung Gear VR[30] are prominent examples of low-cost VR experiences which could provide a solution to the price barrier for mainstream audiences. Consumer smartphones drive the VR experience using a basic, inexpensive headset which, with increased consumer adoption and interest, may lower the price of the entire range of VR systems. While the level of immersion on these devices may not be comparable to high-end systems, the ubiquity of smartphones means that an overwhelming majority of the population potentially has access to a portable device which supports VR.

[29]https://vr.google.com/cardboard/
[30]www.samsung.com/global/galaxy/gear-vr/

The Google Daydream platform mandates specific design and functional requirements such as maintaining head tracking and consistently high frame rates for high-quality mobile VR experiences[31]. However, only the most modern, high-end smartphones can handle Daydream specifications. To achieve a stable, high frame rate, the mobile requires powerful processing chips driving mobile displays with resolutions similar to that of modern desktop computers. This requires expensive hardware, and places a further strain on mobile battery life. However, these limitations will begin to disappear as new hardware emerges and consumers upgrade to the latest generation technology.

Consequently, the opportunities for high-quality mobile VR experiences are set to rapidly increase within the next few years. The Google Cardboard platform has shown that there is an appetite for low-cost mobile VR experiences, with over 5 million headsets shipped over the first 18 months of launching (Bavor, 2016). The pervasiveness of smartphones enables a single device to potentially act as a medium for both self-tracking, and the exploration of personal data in VR. This is only enhanced by smartwatches and other self-trackers which supplement the collection of personal data. Rather than using expensive, high-end VR systems, the progressing specifications of smartphones allow consumers to use their existing device as a means for VR.

2.4 Chapter Summary

This project seeks to evaluate whether VR is a suitable medium for personal data exploration. This literature review has explored a small amount of literature relating to the Quantified Self, the extensively researched field of scientific data visualisation, and its applications in VR. However, a research gap exists between personal data and VR. Could the *Be The Data* visualisation technique be an effective method for personal data exploration in VR? Can the *Parallel Planes* visualisation translate from scientific domain experts to non-expert users in VR?

We have identified a user-driven and system-driven approach to the stage-based model of personal informatics (Li et al., 2010). We have also identified overlapping motivations for initiating self-tracking, and discussed studies which looked at participants goals for collecting personal data (Epstein et al., 2014). The abandonment and barriers of personal data collection were explored, and the challenges of inferring insights from data aggregation platforms were highlighted by Jones and Kelly (2016).

Various approaches towards communicating data to users including dashboards and natural language summaries have been examined through research by Bentley et al. (2013) and as part of a similar study by Epstein et al. (2014). The complexities of filtering multi-faceted personal datasets for visualisation has been studied by Jones and Kelly (2016), and using immersive CAVE-like systems to visualise high-dimensional scientific data has also been described. Furthermore, two visualisation techniques – *Parallel Planes* and *Be The Data* – have been identified to evaluate in the next stages of this project.

[31]https://developers.google.com/vr/distribute/daydream/app-quality

Virtual Reality has been established as an immersive platform which has numerous applications across scientific research, healthcare, entertainment and education. The exhibiting characteristics of immersion and presence have been described in relation to VR (Van Dam et al., 2000; Sheridan, 1992; Slater and Wilbur, 1997). Finally, the ubiquity of smartphone adoption has been pinpointed as a reason which could enable low-cost VR experiences in the future. Evaluating the crossover between Virtual Reality and personal data exploration, and the effectiveness of specific visualisations for non-expert users, forms the next stage of this project.

Chapter 3

Requirements

This chapter begins with an overview of the technology available within the mobile Virtual Reality domain, where we choose Google Daydream as the platform for exploration of personal data in VR. Following this, the implementation path for building a Daydream application is detailed, and requirements are gathered from a multitude of sources, including visualisation techniques discussed during the previous chapter. Given the exploratory nature of this project, these requirements are neither definitive, nor exhaustive. Rather, they serve as a starting point for refining the prototype VR environments during the following Design chapter. A discussion of key requirements elicited from multiple sources concludes this chapter.

3.1 VR Platform Design Space Exploration

During the literature review, we started to collate the immersive characteristics of VR technology. We established that systems must exhibit inclusive, extensive, surrounding, and vivid immersive qualities to generate presence, *the sense of being there*, for the user. We related each of these qualities to certain technological properties. For instance, head tracking with inclusive immersion, and screen richness with vivid immersion. The underlying takeaway was one of a *multi-sensory experience*. This is the idea that VR should touch upon as many of the users' senses as possible.

Clearly to fully meet the objective of a multi-sensory experience we would use standalone head-mounted displays which create the greatest levels of immersion. However, a key idea in this project is the notion of using a single device for collection of self-tracked data, and as the exploratory medium for reflecting upon this data. As a result, this rules out standalone head-mounted displays, and narrows the focus to consumer smartphone VR instead.

VR is an emerging platform, and even more so in the smartphone VR domain. Consequently, there are a limited number of point systems in this space. In fact, there are only three: Google Cardboard, Google Daydream, and Samsung Gear VR. The common factor between these three platforms is that they all run on the Android operating system. Google Cardboard is the only cross-platform VR platform which runs on iOS. It is worth noting that there are various third-party headsets, but these are usually headsets for the Google Cardboard platform, and not a platform in their own right. Consequently, these headsets do not form part of this platform-based discussion.

Android VR Platforms

Given the increased adoption of VR on Android over iOS, plus the authors' existing knowledge of Android, the Android operating system was chosen for development of the VR prototype. Choosing the development platform, however, required a little more considered thought.

Google Cardboard was the first VR platform to arrive on the Android OS. It supports the largest number of Android devices, from Android 4.1 and up giving a potential 97.5% coverage of global Android devices[1]. While the Cardboard platform does have potential widespread coverage, the VR experience is somewhat hindered by platform restrictions to support lower-end devices. The extent of CPU and GPU processing can cause poor responsiveness on low-end devices, adversely affecting users' experiences in VR.

Samsung's Gear VR platform builds upon the basic VR experience offered by Google Cardboard, and adds additional tracking sensors into the Gear VR headset. Combined with low-level GPU optimisations, frame-rate and responsiveness is improved over Google Cardboard. However, Gear VR only supports Samsung smartphones, restricting the audience reach of the VR platform. Additionally, like Cardboard, the only method of interaction is through head gaze[2] which detects where the user is looking within the VR environment and triggers interactions based on their forward facing direction. For rich, multimodal interaction a controller-based approach is preferable.

The Google Daydream platform launched in November 2016 and it is only supported on Android 7.0 and up compatible phones. The requirement for high-end phones with powerful CPU/GPU chips means that there is a reduced audience reach within the current consumer market for this technology. However, this requirement vastly improves VR experiences with compelling visual outputs particularly suitable for the demands of data visualisation. Furthermore, greater requirements are mandated for the design, performance and functionality of Daydream apps – significantly, it is the only mobile VR platform with native controller support. The composition of these immersive attributes, coupled with the additional sensory input through the Daydream controller, led to its selection as the platform to develop for.

[1] Correct as of 31st December 2016. See the Android 'Platform Versions' dashboard: https://developer.android.com/about/dashboards/index.html

[2] In March 2017 after development was complete, Samsung updated the Gear VR platform to include a handheld controller.

3.1.1 Technological Requirements – Building for Google Daydream

At the time of requirement gathering, Google Daydream was at an early developer preview stage. This meant that the Daydream platform was not fully developed, and official/non-official technical documentation was somewhat limited in quantity. However, after a comprehensive review of the features and support available we were able to ascertain that the project was technically feasible. Accordingly, we looked at the 3 paths available for writing a Google Daydream application:

- A C/C++ application with the Google VR NDK (Native Development Kit)
- Unity with the Google VR SDK
- Unreal Engine with the Google VR SDK

Each development path was considered in terms of available resources, and the necessity to rapidly develop prototype applications. The VR NDK was ruled out early on, due to the additional complexity involved with developing against this API. Additionally, we had limited experience with the C/C++ programming languages. The Unity game engine was then chosen over Unreal Engine for 2 key reasons. Firstly, whilst the documentation and support for Google Daydream was sparse – the VR platform had only been announced in recent months – an active community for Daydream developers exists on the Unity forums[3]. If development assistance for this new platform was required, these forums could be used for support. Secondly, we had extensive experience in C# which is the primary language used by the Unity game engine.

A technical preview of Unity (*Unity 5.4.2f-GVR13*) was required to develop for Google Daydream. Although Unity later added native support for Daydream starting with *Unity 5.6*, this had neither been announced nor released at the time of development. Consequently the integration route is through the Daydream SDK, and the manual process of importing SDK *Prefabs* and attaching SDK scripts to Unity *GameObjects*. The Daydream SDK enables the essential building blocks for VR – head-tracking, and stereo rendering with lens distortion. Furthermore, it also contains scripts for controller support and raycasters, both of which are necessary for user interaction. Many of these scripts are simple starting points, and as discussed in the following chapter, some required considerable modification to meet the requirements of this project.

3.2 Requirements Elicitation

The exploratory nature of this project meant that a user-centered approach to determining the full set of requirements was used. Prototyping was used to communicate initial ideas to users, and their feedback influenced purposefully vague initial requirements. The initial

[3]https://forum.unity3d.com/forums/daydream-preview.116/

requirements was based upon analysis of existing data visualisation systems, and Google Daydream guidelines. The iterative prototyping process clarified the design direction of our prototypes by scoping and refining requirements appropriately within the time period available.

The initial requirements was split into functional and non-functional requirements, and labelled as *(FR-1, FR-2 ... FR-N)* and *(NFR-1, NFR-2 ... NFR-N)* respectively. Each requirement was assigned a priority through the MoSCoW technique: *must, should* and *could*. An extensive contract-style list of requirements was not appropriate at this stage as the exploratory system was not rigidly defined. This is in comparison to a traditional software engineering project where system components are tightly specified and interaction between components are strictly controlled. As a result, the requirements were kept relatively concise and flexible, and the use of dependencies and success criteria was avoided.

A discussion follows on the initial requirement gathering within 3 areas: Platform Requirements, Existing VR System Requirements, and Visualisation Requirements.

3.2.1 Platform Requirements

Requirements analysis began at a platform level. The Google Daydream publishing requirements[4] were almost exclusively used to populate the initial requirements in this section. The official Daydream requirements are split into 4 sections: Design, Functionality, Performance & Stability, and Publishing requirements. The Publishing requirements section was dropped from our requirements analysis, as the app was not going to be publicly distributed through the Google Play store[5] and hence the requirements specified in this section were not applicable.

22 requirements were determined at a Platform level from the remaining 3 sections. A full list of platform requirements is available in Appendix A.1. Notable requirements include FR-2: *"The app **must** maintain head tracking"* and FR-16: *"The app **must** maintain high performance and **should** not suffer from thermal throttling"*.

Requirements were either adopted directly from the Daydream developer website, or adapted from the website to fit the objectives of this project. For instance, FR-16 has a corresponding non-functional requirement NFR-16: *"The app **must** maintain high performance across both Be The Data and Parallel Planes visualisations for at least 15 minutes. High performance is defined as a consistently high frame-rate (60fps)"*. This requirement applies the 60fps recommendation for VR platforms which was established earlier in Chapter 2.

Finally, for this section we also looked at specific requirements for the Google VR SDK, Unity, and identified the minimum Android version which this prototype will support. These are given in requirements NFR-1, NFR-2, and NFR-3 respectively. NFR-3 comes with the caveat: *"Daydream is only officially supported on specific Android 7.0 devices"*. We will expand on this further during section 4.2 when we discuss our prototyping environment.

[4]https://developers.google.com/vr/distribute/daydream/app-quality
[5]https://play.google.com/store

This section contributed 18 functional requirements and 4 non-functional requirements. These have largely been based on software characteristics, which stem from hardware specifications for powerful graphics processing and the inclusion of the Daydream controller. The following section now looks at examples of existing data visualisation systems in VR to elucidate further requirements. We will then conclude our requirements analysis by exploring the requirements specific to our two chosen visualisations.

3.2.2 Requirements From Existing VR Data Visualisation Systems

In contrast to the previous section, the requirements identified in this section derive from real-world implementations of VR applications. Given that there are no examples of what a personal data visualisation may look like in VR, we look towards other forms of systems for inspiration on data visualisations in immersive environments. The generalisability of requirements gained from systems which only support non-domain experts may be problematic. Consequently, in this section we will only consider the characteristics of systems directed towards the general population. In the following section we then consider a domain-specific VR data visualisation tool to ensure that we are covering a wide range of potential functionality. A fully fledged prototyping process will ensure that the functionality discovered in this section, and indeed in the following section, is appropriate for the developed system and its audience.

The NASDAQ Rollercoaster

The NASDAQ rollercoaster[6] is the first existing system which we examined for suitable design features (see Figure 3.1). Notable for being one of the first prominent examples of VR used in the news, it is an award-winning visualisation in which the user 'rides' the NASDAQ stock chart over a 20 year period.

Similar to our proposed Be The Data visualisation, the NASDAQ rollercoaster immerses the user right within the data. The user is placed at the top of the chart and moves forward as time progresses. This experience is supplemented by storytelling throughout, providing additional contextual information around sudden peaks and troughs in the visualisation. The user is able to look around 360°in all angles. The axes design and explanatory text is well considered – in particular, the axes labels are reduced only to the most significant data points. This reduces an overload of information, keeping the visualisation simple and increases the impact of the data itself.

This led to the identification of 9 functional and 2 non-functional requirements. Notable requirements include FN-20: *"There **must** be a smooth transition when moving from an overview of the data to a detailed point inside it"*. This has a corresponding non-functional requirement NFR-6, which specifies that movement must follow a linear or bezier curve between points. Additionally, FN-26 states that *"Textual information **should** be avoided*

[6] http://graphics.wsj.com/3d-nasdaq/

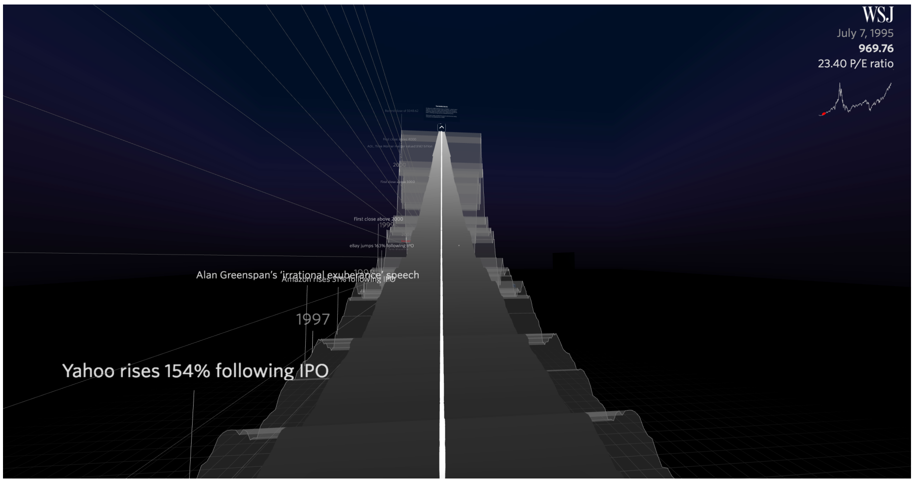


Figure 3.1: Screenshot of "Is the Nasdaq in Another Bubble? A virtual reality guided tour of 21 years of the Nasdaq" (Kenny and Asnes Becker, 2015)

where possible. Let the data do the speaking.". This requirement stems from the reduced textual information in the NASDAQ visualisation, and indeed in Chapter 2 where we found that visual representations of data provide a "*higher degree of confidence in the findings*" than textual information does (Keim, 2001, p.40). However, this was for scientific specialists looking at large datasets. Therefore, at the design stage in our system we will initially give preference to the data over textual representations, and verify that users do not need extra textual information during the prototyping process.

Limitations of the NASDAQ visualisation also influenced our requirements. The objective of the NASDAQ visualisation was to follow a rollercoaster-like path on top of the data chart. However, this was a predetermined path and there was no ability to move off it and view the data from a different angle in the environment. Accordingly, FN-27 was devised, targeting the free movement of the user in the data exploration environment. Furthermore, the NASDAQ visualisation did not include sound which was discussed in Chapter 2 as an important factor for presence. Correspondingly, sound effects through FN-28 was established as the final requirement emanating from this existing system.

Will the UK Brexit?

We then used a secondary VR system to validate our requirements. The target system was named 'Will the UK Brexit?'[7], created by data journalists at Google News Lab (see Figure 3.2). The visualisation allows users to highlight European Union countries and see their population's most searched Brexit-related questions in the runup to the UK's EU membership referendum in 2016.

[7]http://news-lab-brexit.appspot.com/vr/

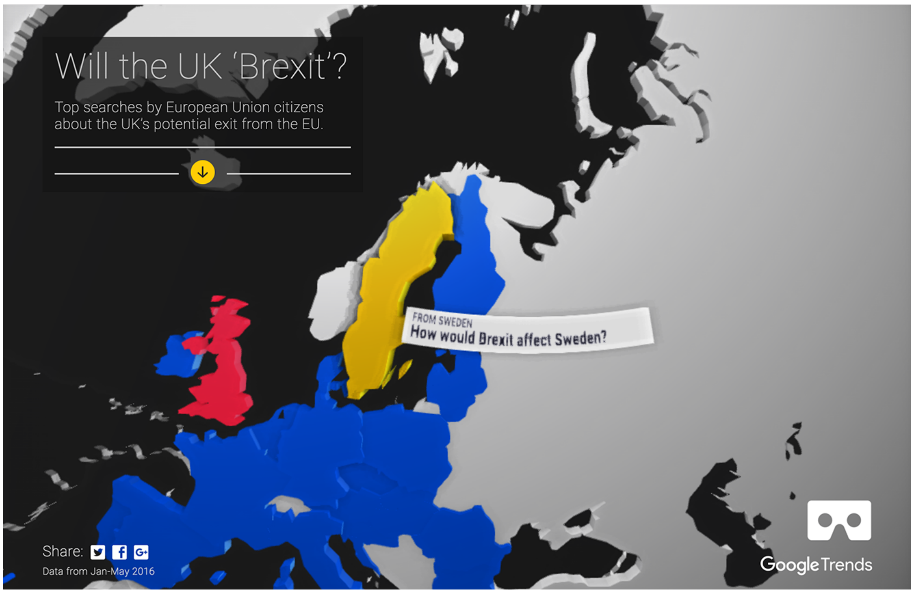


Figure 3.2: Screenshot of "Will the UK Brexit?" (Lab, 2016)

Again, the VR data visualisation contains little textual representations. Popular search engine queries appear for each country only when the user has highlighted that specific country. This further substantiates our findings for FN-26 around reducing the amount of text where possible. A new set of requirements {FN-29, FN-30, NFR-7} extend this data highlighting mechanism, adding the concept of a highlight colour and a data overlay containing additional information about the data point (as seen in Figure 3.2). NFR-7 specifies FN-30 in that the *"information overlay **should** fit entirely inside the VR stereo rendering"*. This requirement arose from our testing of this visualisation, where to fully see certain search queries, we had to move our head away from the selected country. Our final requirement for this section was inspired by the instructions which this visualisation used should the user look behind them. These instructions point the user back towards the main area of interest – the data visualisation. FN-31 therefore defines a low priority requirement to use visual cues should the user stray too far from the data visualisation.

Two visualisations – 'The NASDAQ Rollercoaster' and 'Will the UK Brexit?' – were chosen for requirements analysis for our system. With a sparse number of data visualisations in VR aimed at the general population, these visualisations both came from the data journalism field, and hence were selected due to their intended wide-reaching audience. The full list of requirements identified in this section is listed in Appendix A.2. Having established platform requirements, and having begun to identify more specific app requirements, the next section considers the two data visualisations we will be implementing – Be The Data and Parallel Planes – in greater detail.

3.2.3 Visualisation Requirements

This section considers the requirements for the two visualisations which we will be implementing. A number of requirements have already been discussed in the previous sections which will serve as the foundation for the requirements explored here. We will also look at the existing VR Parallel Planes literature and determine the functionality appropriate to non-experts. Prototyping will enable these initial requirements to evolve to a scope suitable for our exploratory study.

Be The Data

The Be The Data study by Chen et al. (2016) was introduced in Chapter 2. In their study, Chen et al. placed students into a physical environment and each student became an individual data point representing a point in a system. Students could then collaborate with their immediate neighbours, and cluster or separate accordingly in response to data exploration questions. Our proposed system differs in several respects to Chen et al.'s study. Firstly, it will not be a collaborative system. Users will explore the data individually, rather than alongside others. Secondly, the data exploration process will take place inside a virtual environment, rather than a physical one. The fundamental idea which we are elaborating on is one of data embodiment – becoming a single data point in a larger collection of data points – and seeing whether this can positively benefit the users' understanding of the dataset.

Chen et al. give no indication on how this might be achieved in a virtual environment. As a result, the requirements identified in this section largely build upon previously established requirements, in the context of exploring a standard three-dimensional scatter plot in VR. To start informing our thinking around prototype requirements for this visualisation, we used Matlab to generate a 3D scatter plot with random data points.

Figure 3.3 shows our randomly generated graph. We identified 7 functional requirements and 4 non-functional requirements by thinking about actions which a user might take when exploring this form of graph in an immersive, VR environment. {FN-32 ... FN-38} and {NFR-8 ... NFR10} are the non-exhaustive list of requirements for this section, with the full list available at Appendix A.3. As we do not know what the needs of a user will be with this visualisation, or their typical exploratory behaviour, these requirements are incomplete estimations which will be validated during the prototyping process.

Nevertheless, this brought about some key requirements. Namely FN-32 and FN-33, which state that the user's movement must not be *"constrained by the graph boundaries"* and that the user must be able to move outside of the graph through the axes. FN-38 extended ideas on movement directions from the previous section, specifying that there should be at least 4 directions that the user can move in. Several requirements relevant to this visualisation were already covered in previous sections, such as FN-28 and FN-30. Consequently, non-functional requirements were detailed to constrain exactly how the system will achieve these initial requirements.

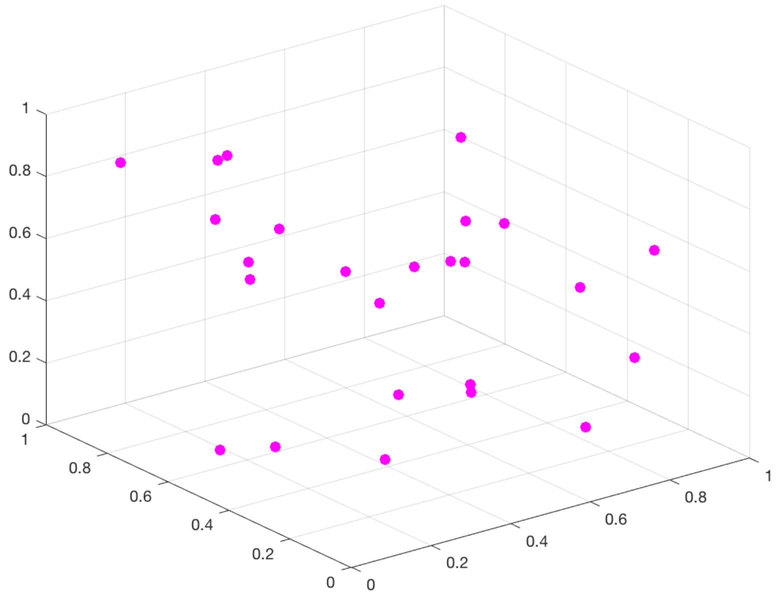


Figure 3.3: A randomly generated 3D scatter plot used to conceive initial requirements

With a small number of requirements evoked from the randomly generated graph, we had an idea of the initial requirements for this visualisation. During the prototyping process, when we get an improved idea of the mental models of users while interacting with this system, we will refine these requirements appropriately.

Parallel Planes

Our second visualisation selected in Chapter 2 was the Parallel Planes visualisation. This extends the traditional parallel coordinates visualisation technique by adding an extra dimension in the z axis. Figure 3.4 shows a visual representation of a data observation through $x_1, x_2 \ldots x_n$. Brunhart-Lupo et al. (2016, p.2) describes this visualisation as a *"series of scatter plots where the same observations are joined by a polyline"*. Two additional features extend these series of scatter plots in their visualisation system - specifically data brushing, enabling observations to be highlighted amongst the entire dataset, and the running of additional simulations on the brushed region. Given that we will only be visualising a pre-determined personal dataset, this latter feature of simulation is not applicable to our system.

However, brushing is a compelling feature which we will be specifying for our Parallel Planes visualisation system. The ability to select a subset of data observations passing through particular regions has clear use cases. Our prototyping process will aim to make

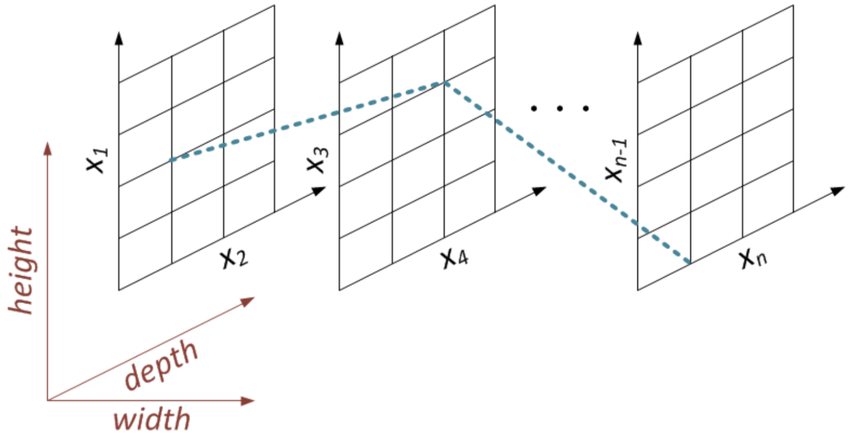


Figure 3.4: Parallel Planes visualisation of $x_1, x_2 \ldots x_n$ (Brunhart-Lupo et al., 2016)

this feature intuitive and rewarding to use. Accordingly, {FN-41, FN-42, NFR-12} were developed as brushing requirements. These requirements specify that the user *must* be able to select *"at least one"* subset of the data, and that they *should* be able to select multiple subsets. Once selected, the non-functional requirement stipulates that the selected subset of lines must have a highlight colour, and that non-selected data observations should be dimmed.

Additional requirements conforming to the structure of the visualisation were also prepared. For example, FN-39 demonstrates that each new dimension of the personal dataset must be represented on a new plane. Notably, this led to the development of a new requirement FN-40, which stipulates that the dimension in the z axis of each plane must be kept constant. With the Be The Data visualisation, we were restricted to displaying 3 dimensions of data in each axis (though this could be extended to 5 dimensions by changing the shape and colour of the data points). In the Parallel Planes visualisation, adding a new dimension is completed by adding a new plane. Potentially this lets us represent n dimensions, though in practice n would be low due to visual and cognitive limitations. Particularly given the target audience of this system is not subject matter experts like in Brunhart-Lupo et al.'s study, when working with a high-dimensionality of data, keeping the z axis a constant dimension (such as the Day of the Week) should help to reduce the typical interpretation challenges associated with multi-dimensional data.

With requirements for movement and sound in the VR environment recorded previously, a total of 5 functional requirements and 1 non-function requirement were collected from this section on Parallel Planes. The full set of initial requirements are available in Appendix A.4.

3.3 Chapter Summary

In this chapter we completed an exploration of the VR platforms suitable for the development of our prototypes. Google Daydream was selected as the target platform, and Unity and the Google SDK were chosen as tools to develop the visualisations with.

We gathered 43 functional requirements and 12 non-functional requirements through an elicitation process covering platform requirements, analysis of existing systems and requirements specific to our two chosen visualisations. These requirements are neither exhaustive, nor extensive, but are initial starting points for refinement during the prototyping process.

The next chapter considers the refinement of these requirements as the design of the visualisations progresses through the prototyping process.

Chapter 4

Design

This chapter focuses on the key design considerations we made when developing the visualisations as part of an iterative prototyping process. We begin by discussing the technique used to pre-process the dataset represented in the final visualisations. The prototyping process is then outlined and we build upon and refine the requirements established in the previous chapter.

Design considerations such as axes representation and interaction in VR are highlighted in terms of both visualisations. Multiple figures of early prototypes and the final prototype are included to illustrate and support the decisions made as part of the prototyping process.

The manual swipe implementation for the Google Daydream controller is detailed, and the chapter concludes with a discussion on design decisions made around realism and missing data.

4.1 Data and Data Preparation

During Chapter 2, we reviewed the rise of personal data tracking and analytics. Through literature by Jones and Kelly (2016), and Choe et al. (2014), we highlighted the cognitive challenges of examining large, correlational datasets. The prototypes we develop over this section use a multidimensional personal dataset, with the dataset visualised within an immersive VR environment. Our insight-based evaluation will assess whether the qualities of VR can reduce the interpretation challenges associated with multi-dimensional data.

Due to time constraints on the project, it was not feasible to collect several months worth of data such that participants could later use their own data in the VR visualisation. This limitation is discussed later in Chapter 8. Instead, Dr Simon Jones provided a dataset containing tracked personal information about 19 participants. This dataset contained up to 36 attributes about each participant tracked over a period of 2-3 months. Examples of participant attributes include *weather, productive minutes, steps,* and *events.*

The dataset we used was provided in the form of an SQL dump. Previously this database had been used on the self-tracking data aggregation platform Exist.io[1] and contained two tables: *attributes* and *correlations*. The correlations table related two attributes together, calculated a correlation strength, and generated natural language statements such as *"You get more steps when you listen to more music"*. However, we were only interested in the raw recorded data observations present in the attributes table. Accordingly, the correlations table was manually discarded, and our focus turned to preparing the attributes table for our prototype.

4.1.1 JSON Serialisation

We found it challenging to connect the Unity application with the local SQL database – there was no native support for this offered by Unity. Instead we looked for alternative solutions for getting the participant data into Unity. Our final technique involved using a JSON format to represent the SQL database, and then using Unity's JSONUtility to serialise the database into C# objects.

The first step of this process involved converting our SQL database into a JSON text-format. With no online tools available to do this job, we defined a JSON schema, and wrote a text parser in Python to translate SQL statements into a JSON instance.

Listing 4.1: JSON Schema

```
{
  "dbName": "<database_name>",
  "participants": [{
      "participantName": "<participant_name>",
      "attributes": [{
          "attributeName": "<attribute_name>",
          "attributeData": [{
              "date": "<date>",
              "value": "<value>"
            },
            {
              "date": "<date>",
              "value": "<value>"
            }
          ]}
      ]
    }
  ]
}
```

[1]https://exist.io/

Listing 4.1 defines the structure of the JSON schema which we used for our database. The main database object contains an array of *participants*. Each participant has a name and an array of tracked*attributes*. Each attribute has a name, and has an array of *attribute data*. These objects contain the daily data observations pertaining to that specific attribute.

The Python parser accepts SQL statements, parses the input and organises the participants and their recorded attribute data into a single JSON file. For instance, given a single participant with two attributes, the Python script outputs the JSON object in Listing 4.3.

Listing 4.2: Example SQL Input

```
('participant1','distracting_min','2015-07-26','55'),
('participant1','distracting_min','2015-07-27','128'),
('participant1','mood','2015-07-26','5')
```

Listing 4.3: Example JSON Output

```
{"dbName": "simonsdata",
  "participants": [
    {
      "participantName": "participant1",
      "attributes": [
        {
          "attributeName": "distracting_min",
          "attributeData": [
            {
              "date": "2015-07-26",
              "value": "55"
            },
            {
              "date": "2015-07-27",
              "value": "128"
            }
          ]
        },
        {
          "attributeName": "mood",
          "attributeData": [
            {
              "date": "2015-07-26",
              "value": "5"
            }
          ]
        }
```

```
        ]
      }
   ]
}
```

The Python script was run over the 35,000+ line SQL database containing 19 participants and their tracked daily data points during a 2-3 month period. Significantly, this is a one-time process, and so will not have any adverse effect on app startup time for the prototypes. The output JSON formatted file contained a list of participants and their associated attributes. With the data now in a format appropriate for use in Unity, Unity's JsonUtility class was used serialise the JSON instance into C# objects.

Listing 4.4: Participant Class

```
[Serializable]
public class Participant{
        public string participantName;
        public List<Attribute> attributes;
}
```

Listing 4.4 shows one of four serialisable classes declared in our Unity code. Finally, the JsonUtility function reads in the precomputed JSON instance at run time, and serialises the data using these classes into C# objects. We are now able to access our data in an object-oriented fashion. For example:

Listing 4.5: Example Data Access

```
// Get the first participant
Participant active = jsonDatabase.participants [0];

// Return the number of productive minutes from their 29th day
    of data logging
return active.attributes [9].attributeData [29].value;
```

In the absence of native local database support, this object-oriented approach lent itself well to the needs of our project. After serialising the dataset, we had the ability to pass around objects within the Unity game engine, enabling the dataset to be adjusted as needed during different runs of the prototype. The cuts of this dataset chosen for use in our visualisations is discussed in greater detail during Chapter 6. A flow chart capturing the pre-processing steps we took on our dataset is detailed in Figure 4.1.

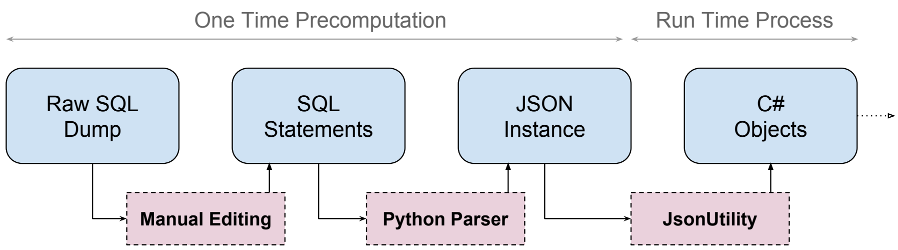


Figure 4.1: The Data Pre-Processing Steps

4.2 Prototyping Process

With our personal dataset now ready to use, we progressed onto designing the visualisations in Unity. Having set out our initial requirements in Chapter 3, we built a basic first prototype to gather feedback on usability and user experience goals. With the time constraints considered, we planned a prototyping process in which we would be able to evaluate the systems twice with test users, before releasing to our study participants. This corresponds to two full cycles of Figure 4.2.

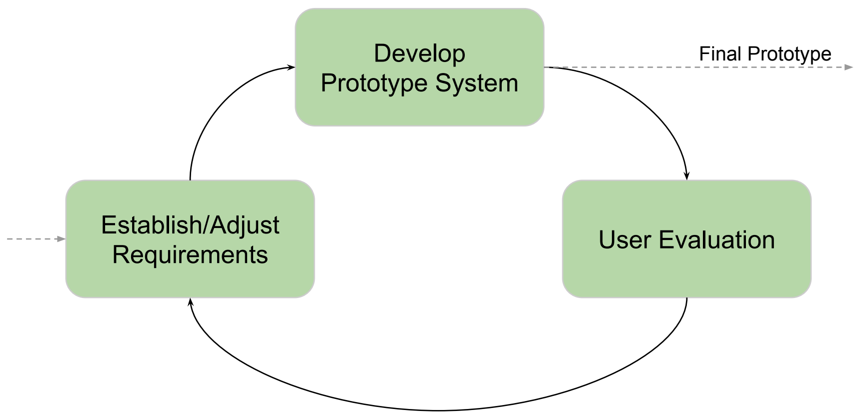


Figure 4.2: The Prototype Process

The cycle began by establishing the initial requirements determined in Chapter 3. A prototype was developed according to this requirements specification, and released to users for evaluation. In our case we used 5 test users for both evaluation phases, with 3 users constant between phases and 2 new users at each phase. These were formative evaluations, but with respect to their structure, quite informal and qualitative in nature. At each phase we evaluated the prototype in an open and undirected way with the user. For instance, we followed users' typical navigation through the data exploration environment, and identified typical tasks and paths which users undertook. Furthermore, we asked questions around the presentation of the visualisations, and ensured that the users were able to get a grasp of the environment in the 10 minute time period we allowed.

The evaluation feedback was then used to define new requirements, or indeed adjust and rescope existing ones. After two phases of this feedback, we built our third and final prototype which we used for our study.

4.2.1 Prototype Environment

Before characterising specific prototype features highlighted through user feedback, we will describe the testing environment which we used during prototyping. In the first phase of the prototyping process, we targeted the application we built at the Google Cardboard platform. This enabled us to develop the prototype visualisations rapidly, and demonstrate a basic working example of our application to users. This prototype covered all of our initial requirements, with the exception of control-based requirements – these requirements had been written with the Daydream controller in mind. Instead, for the first prototype, interaction was initially completed through head gaze and a small capacitive button on the Cardboard headset.

For the second and third prototype, we switched our target platform to Google Daydream. This allowed us to add Daydream controller support, enabling users to point at objects in the virtual environment with their hands, rather than searching through head gaze. In our first phase prototype, we had already completed the groundwork for movement and interaction with our Cardboard app. Building a Daydream app in the second and third phases simply involved adding Daydream controller support on top of this integration.

Our testing smartphone across all phases was a Nexus 6P running Android 7.0 Nougat. When we began to target the prototype at Google Daydream, an additional app-based modification was required. This involved sideloading the Google Daydream Home APK[2] and selecting *Skip Entry VR Screens* in the app settings. This enabled apps like our prototype which target the Google Daydream VR platform to run on the Nexus 6P.

Officially the Nexus 6P does not support Google Daydream, although this phone was the recommended development device for the platform before the release of officially supported Daydream phones. The Daydream developer setup page[3] comes with the caveat that the Nexus 6P's "*thermal performance is not representative of the consumer Daydream-ready devices*", which may lead to CPU and GPU throttling. However, in practice we found that our visualisations were not technically demanding enough to have any adverse effect on performance.

Having discussed the prototyping process, the following sections describe how we went

[2]https://play.google.com/store/apps/details?id=com.google.android.vr.home
[3]https://developers.google.com/vr/daydream/dev-kit-setup

about addressing specific design challenges of visualising personal data in VR.

4.2.2 Axes and Label Representation

Feedback relating to the design and representation of axes and data labels was undoubtedly the most numerous during our prototyping phases. Over this section we will cover the various strategies we used to improve the axes in both of our visualisations, such that users could read and interpret our visualisations clearly.

Be The Data

In our first prototype we included 3 partially transparent planes representing the axes along the X, Y and Z components of the dataset. These 3 planes can be seen in Figure 4.3.

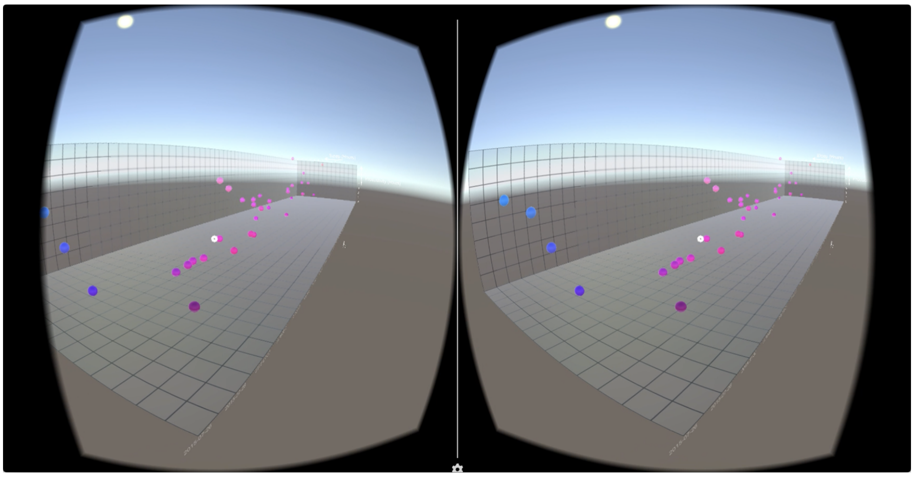


Figure 4.3: Early Prototype: 3 partially transparent planes *(Stereoscopic View)*

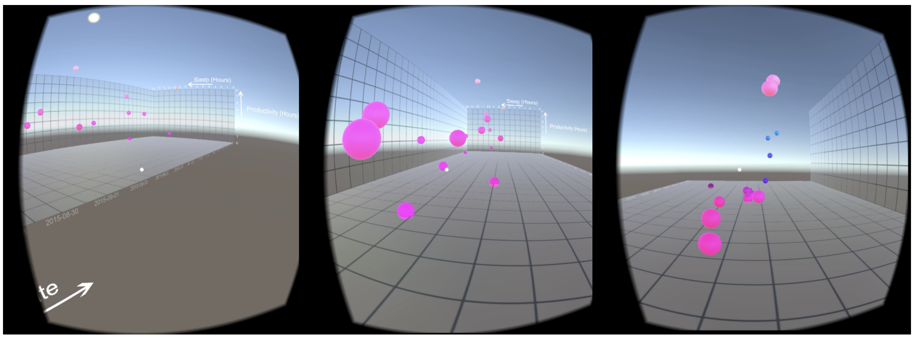


Figure 4.4: Early Prototype: Typical exploration path series *(3 Separate Left-Eye Views)*

During the first phase of prototyping, we discovered that users generally followed a path from their opening perspective, in which the axes played a crucial part. Figure 4.4 shows a series of 3 images representing the typical path in which a user initally navigated though the data. The leftmost image shows the opening position in which a user is placed on entering the Be The Data visualisation. This is half-way down the X axis, and slightly raised up looking down over the graph. This enables users to obtain an overview of the entire graph and identify the major axis dimensions by looking to their right. The second image in Figure 4.4 shows the position of the user after clicking on a data point in the centre of the graph. Selecting a data point first, rather than moving freely around the graph, was generally standard behaviour for our prototype participants.

At this point in the visualisation, the user is immersed in the data and can look down the X axis in either direction, potentially to try and understand correlations in the data from before and after their position in the environment. However, this is where our first prototype became problematic. The second image in this series shows that a plane representing the Y and Z axes is present at the $X = \max(\mathrm{x})$ position, but not in the third image where $X = \min(\mathrm{x})$. From the evaluation feedback it was clear that two planes were required in both positions, so that users could make 2D projections of the 3D data against each plane. Consequently we added an additional plane at $X = \min(\mathrm{x})$ so that users could look both up and down the X axis. Screenshots of this are available under Implementation Results (Chapter 5).

The next design change we made as a result of prototype feedback was adding arrows to illustrate the direction of magnitude of the axes labels in each dimension. This stemmed from users being positioned at one end of the X axis and looking back down the axis. The labels on the plane representing the Y&Z axis at the other end were often not clear enough, purely due to the distance between the label and the user. While the user had the ability to turn around 180 degrees and see the same axis at the end they were currently positioned at, this exerted additional physical effort and switching from one visual context to another was complex for participants. These arrows therefore let users understand general trends over the breadth of the dataset, rather than a deep exploration of a single data point. Figure 4.5 shows how these arrows appeared alongside the axes in our final prototype.

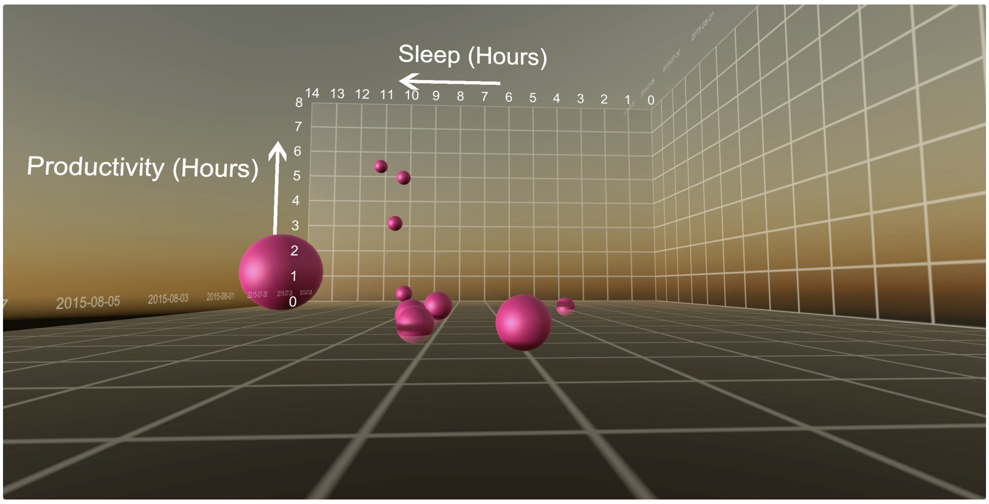


Figure 4.5: Final Prototype: Arrows in the direction of label magnitude

A challenge unique to this data exploration environment was that data labels could not remain static as the user moved around in the environment. In particular, during our prototyping sessions users both looked over the data (Figure 4.6) and also immersed themselves inside the dataset (Figure 4.7). In our first prototype, the data labels remained in the same position regardless where the users placed themselves – either outside of the graph, or inside of it. In the example screenshots, this meant that if the user looked back to the X axis to read the date (as in Figure 4.7), the labels were out of view below the plane, and also the wrong way around. In our final prototype, we made the labels dynamic so that they responded to face the user's position in the environment. Figure 4.6 and 4.7 shows our solution for changing the position and rotation for all of the data labels along the X axis.

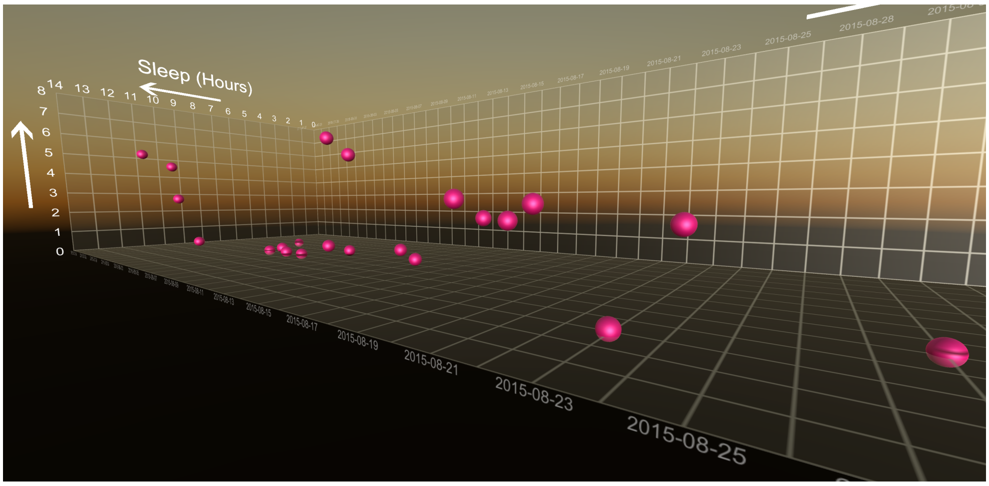


Figure 4.6: Final Prototype: Dynamic Labels

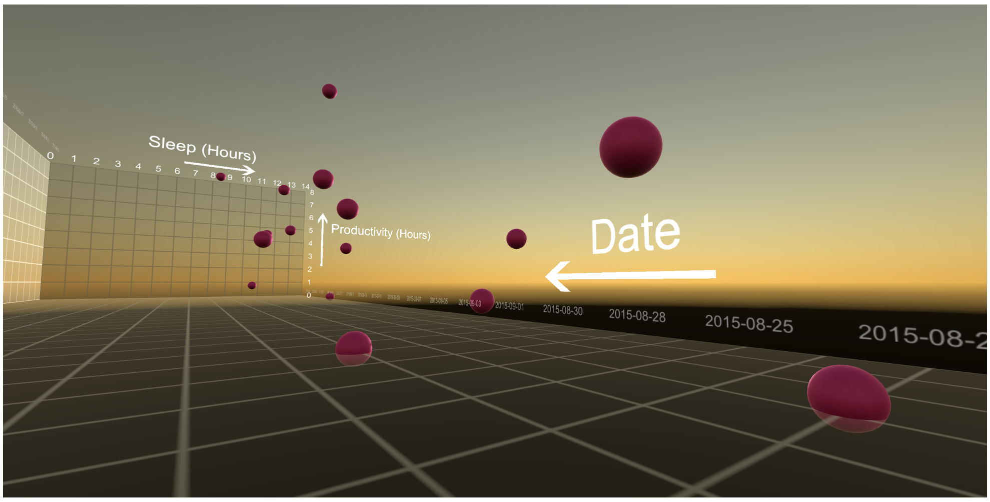


Figure 4.7: Final Prototype: Dynamic Labels

Listing 4.6: C# method for label rotation

```
// Rotate X label (e.g when user crosses the Z=0 plane)
// Shift the label up and below the axis depending on position
public void rotateXLabel(float yRotation, float yShift){
        this.gameObject.transform.GetChild(0).
        gameObject.transform.rotation = Quaternion.Euler(new
            Vector3(0f, yRotation, 0f));

        Vector3 newPos = new Vector3(
        this.gameObject.transform.GetChild(0).
        gameObject.transform.position.x,
        yShift,
        this.gameObject.transform.GetChild
            (0).gameObject.transform.position.z);

        this.gameObject.transform.GetChild
            (0).transform.position = newPos;
}
```

When a change in the user's Z position is detected involving a move across the $Z = 0$ plane, the method in Listing 4.6 is called with suitable parameters and iteratively applied to a data structure containing all the environment X labels.

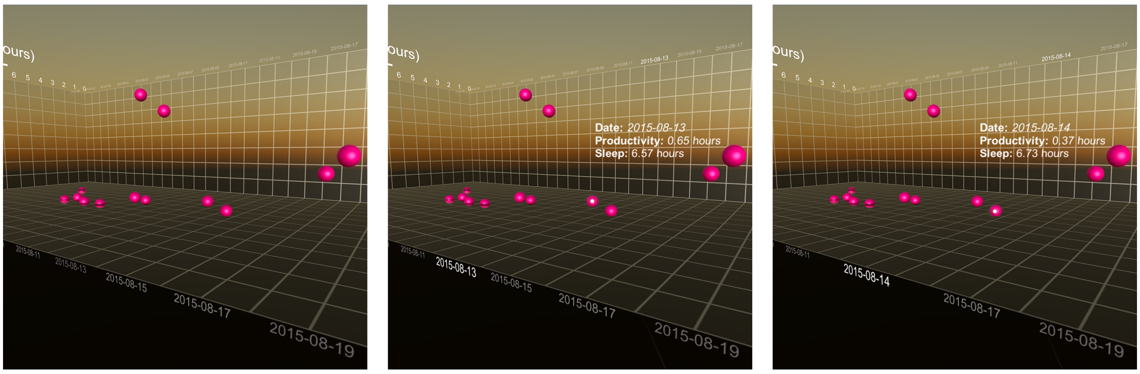


Figure 4.8: Final Prototype: Selection Emphasis *(Non-VR View)*

Left: No selection. Centre: Selection on visible label. Right: Selection on a previously hidden label.

The final change we made to the Be The Data visualisation came after some participants spoke about the difficulty of locating the position of a selected data point in the X axis when viewed from certain angles. While the textual overlay which appears on selection was useful in providing contextual information about the data point, what participants wanted was a form of visual indication of where the data point sat along the X axis. To achieve this, we emphasised the X axis label corresponding to the X dimension of the user selected data point. A scaling animation grew the data label to a larger size, and switched the label from transparent to opaque. Data labels surrounding the selected data label were also hidden for added emphasis.

This prototype feedback did have a small conflict with requirement FN-24 which we had implemented already. This requirement specified that "*the number of axes labels **should** be reduced where possible*" and we were meeting this requirement by hiding every other axis label. Therefore, we devised a solution which was able to highlight both hidden labels, as well as already visible labels. Figure 4.8 shows the results of the user selecting a data point which has a visible data label (*centre image*), and a different data point which has an initially hidden label (*right image*). Ultimately this reduced inaccurate interpretations of axis readings due to the angle the user was positioned at, and supported users in spatially locating the data point in the data exploration environment.

Parallel Planes

Prototyping raised fewer axis-related points in the Parallel Planes visualisation, generally because users did not move around in the environment as much compared to the Be The Data visualisation. Nevertheless, there were two design changes which we made for this visualisation as a result of prototype feedback.

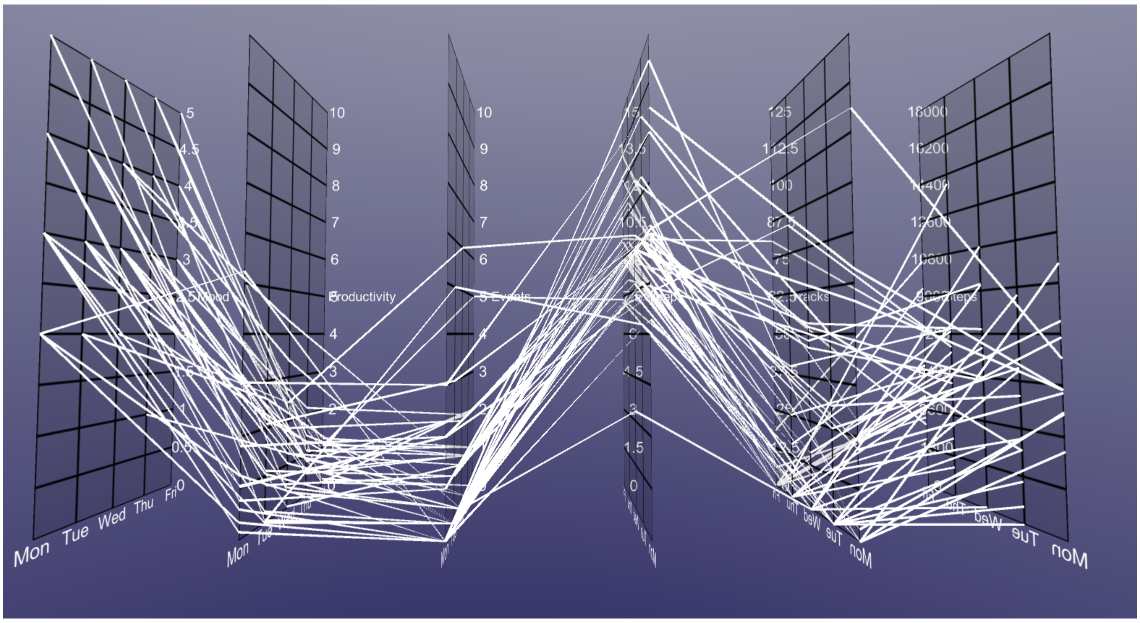


Figure 4.9: Early Prototype: Axes Incorrectly Scaled

The first of these changes relates to the maximum unit on each plane, and is a matter of normalising the data on the plane appropriately. For example in Figure 4.9, the second plane shows the Day of the Week in the X axis, against the number of productive hours in the Y axis. The intersection of a line on the plane is the number of productive hours on a certain day. In the dataset used on this early prototype, there was not a day where the number of productive hours was greater than 6 hours. However, this plane has a maximum unit of 10 productive hours, which results in the lines being tightly clustered at the bottom of the plane. Prototyping feedback suggested that it was complex for users to discern patterns in the dataset when the visualisation did not make use of the entire plane to spread out the dataset.

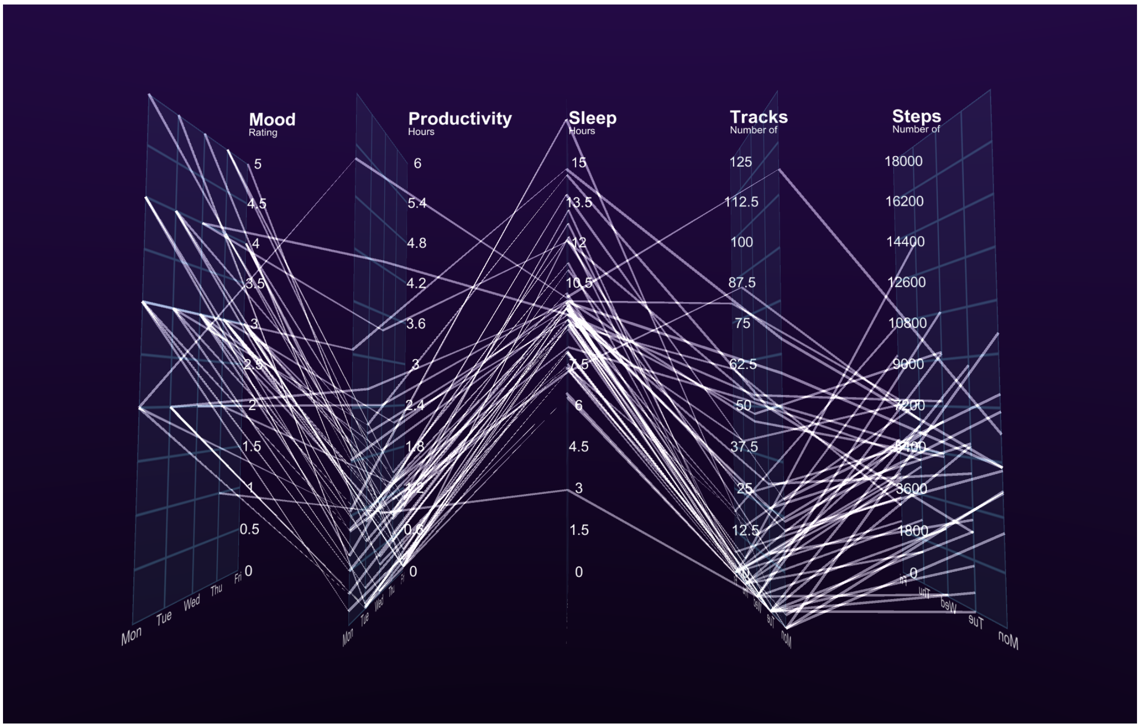


Figure 4.10: Final Prototype: Axes Correctly Scaled

Figure 4.10 shows the results of implementing user feedback on scaling the data appropriately to the planes. Compared to Figure 4.9, the lines intersecting with the productivity plane make better use of the space available (*Note: The 3rd plane from Figure 4.9 has been removed*). It is now easier to comprehend which lines are outliers to the main cluster of productivity due to the increased distance between distinct lines. This does raise a further question of how best to deal with significant outliers which impact the scaling determined on each axis, and lead to even tighter clusters of lines. One large outlier can completely change the complexion of the visualisation. This is not just exclusive to our Parallel Planes visualisation in VR – it also affects traditional 2D Parallel Coordinates visualisations as well. Suggested future work relating to this design challenge is discussed in Chapter 8.

The second design change we made as a result of prototyping feedback was to increase the transparency of the planes. While it was not general behaviour to look down the planes in a side-on view, a few participants did interpret the data from this new perspective. However, the transparency of the axes prevented the observation of any meaningful patterns in the data – successive plane backgrounds built up to build a single opaque plane in which the lines could not be distinguished against. Consequently, as the number of dimensions (planes) grows in the Parallel Planes visualisation, this angle of observing data could become problematic. A temporary solution was put in place which enabled participants to look down the planes and distinguish the lines against all the planes (Figure 4.11). Nevertheless, should the number of dimensions rapidly increase a different solution would be necessary, should this viewing angle be required.

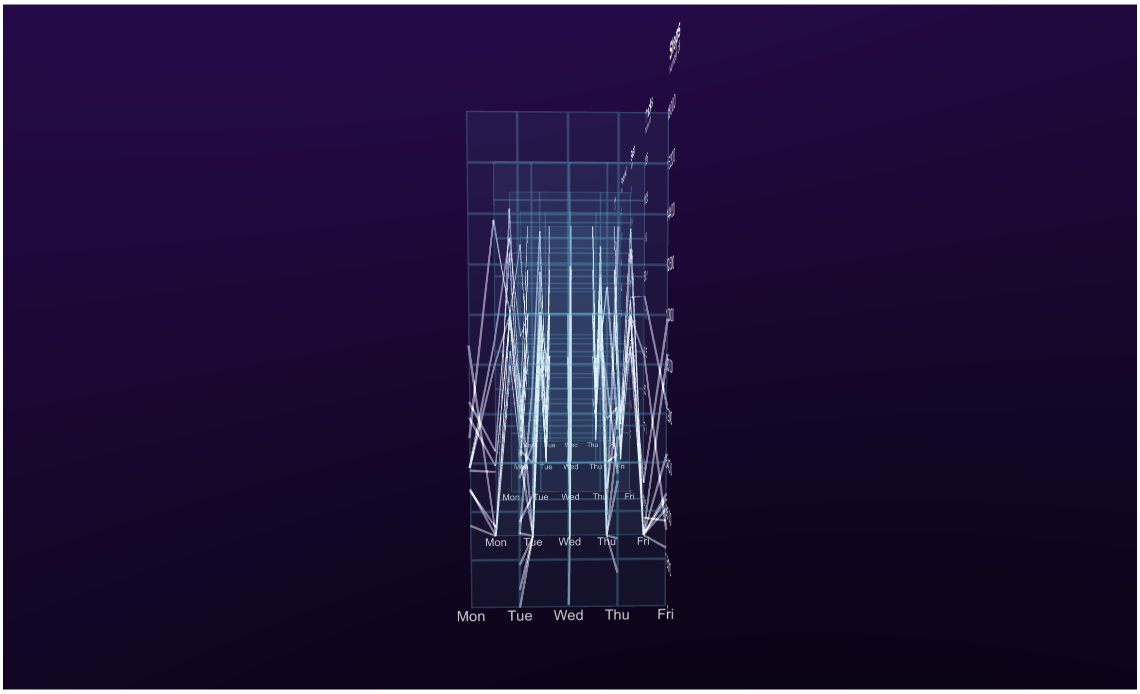


Figure 4.11: Final Prototype: Side-on view

4.2.3 Interaction: Movement and Selection

This section looks at the implementation of interactive controls in the visualisation environments. The development of our movement functionality will be explained in this section, followed by a description of interaction with components of the data visualisations.

The earliest prototype was built for the Google Cardboard platform, and for this reason, interaction in this environment was restricted to the basic head gaze technique. With this technique, a small cursor named a 'reticle' sits in the centre of the user's view, and expands into a circle when the user looks over an object which has interactive properties. In our Be The Data visualisation, prototype participants followed a typical path from their initial starting point. This corresponds to the images from left to right in Figure 4.12.

- 1: Initial starting sweep – looking around for data points to interact with.
- 2: Identify target – reticle expands into a circle.
- 3: Click target – user accelerates towards the target
- 4: Look around from target – user positioned at the target

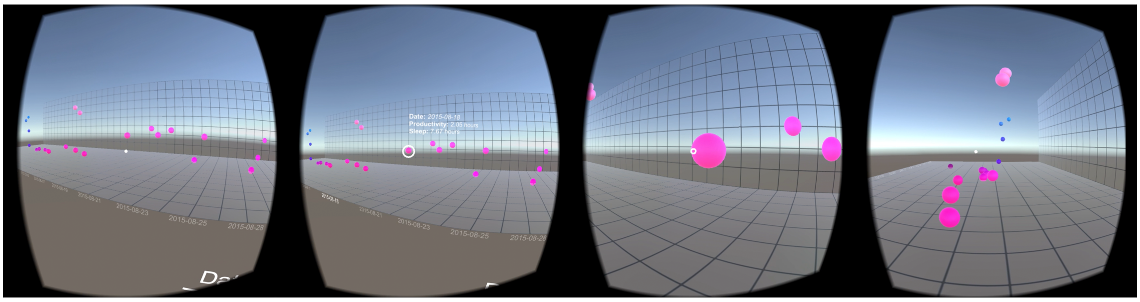


Figure 4.12: Early Prototype: Head-Gaze Interaction Series

While this provided a quick way for the user to get immersed in the data, head gaze came with several limitations. The user could only move between data points with this method – they were not able to move freely in the world and view the data from different angles. Thus, their movement in the environment was predetermined by the structure of the data. Additionally, the 'click' was through a capacitive button on the Cardboard headset which was not natural. The user had to reach to the headset every time they wanted to move.

Switching to the Daydream platform enabled us to use the more comfortable Daydream headset and, significantly, the Daydream controller. The benefit of the controller is that it is entirely separate to head movement. It appears independently in the environment (as in Figure 4.13), acting as a laser pointer at objects. When the user moves the controller in the real environment, the controllers movement is mimicked in the virtual environment. Buttons on the device direct interaction with objects in the data visualisation. Several prototyping participants who had experienced head gaze interaction previously commented on how much the controller improved the VR experience, particularly for making the data selection process "*more natural*".

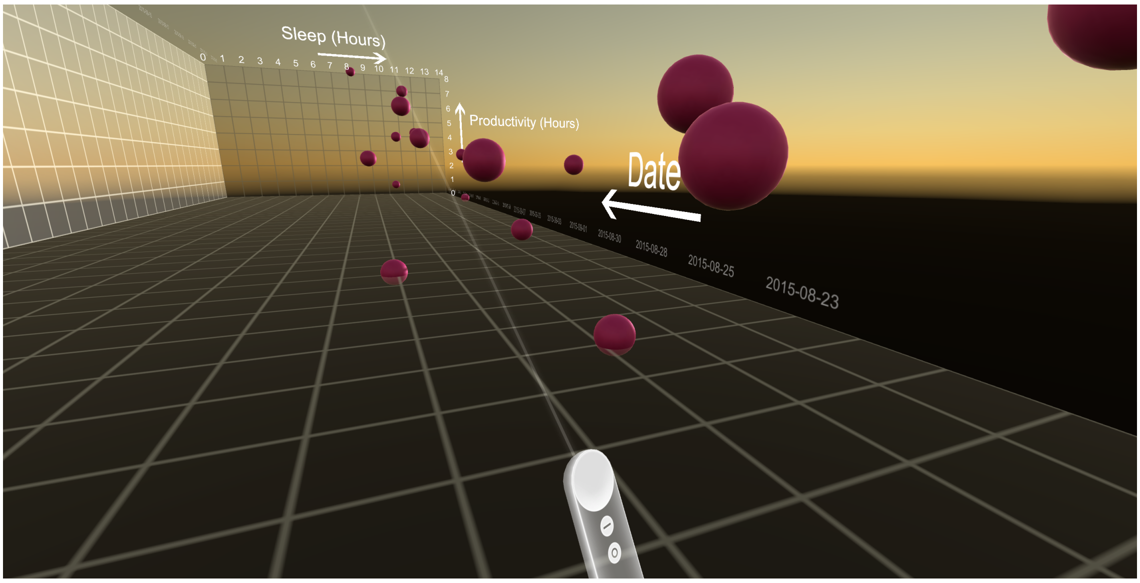


Figure 4.13: Final Prototype: Daydream Controller Laser

Figure 4.13 shows the Daydream controller raised up into the user's view. The controller contains two small buttons (App and Home buttons) which we did not use in our prototype. We did however make extensive use of the circular trackpad at the top. This trackpad detects the position of a user's touch, and can also be pressed inwards as a click. This enabled users to aim at a data point using the controller, and then click on the touchpad to move towards the target in the virtual world. Users could also swipe on the trackpad to move forward and backwards, and strafe from side to side. Swiping was not built into the Google SDK, and so this feature had to be implemented manually.

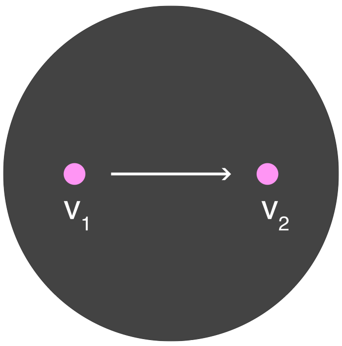


Figure 4.14: Daydream Controller: Swipe Implementation

The Google SDK returned a simple two-dimensional vector of the user's latest touch position on the trackpad. We recorded the user's first touch position on the trackpad, and the final position on the trackpad before their finger was lifted off. These positions correspond to v_1 and v_2 in Figure 4.14. We also recorded the elapsed time between v_1 and v_2 occurring. This enabled us to calculate the velocity of the swipe between touch points, and adjust the users position in the environment according to the speed of this swipe. Velocity was calculated as:

$$\frac{v_1 - v_2}{\Delta v_1 v_2}. \tag{4.1}$$

A low-pass filter was sourced from the Google SDK to improve the accuracy of the velocity, and this was applied with Unity's *Vector2.Lerp* method to linearly interpolate to the target velocity vector. Finally, if $\Delta v_1 v_2 > kClickThreshold$ – a predetermined threshold to prevent the user moving when pressing the touchpad inwards as a click – the velocity is sent to our method *daydreamWorldPositionHandler()* to determine the user's new position. This is based on the direction of the swipe (left, right, up, down), the velocity of the swipe, and the user's forward facing direction.

With the trackpad swipe giving a final position in the world through the calculations above, we ensured that the user's movement between the start and end positions was realistic. We could not simply update the user's position between two frames – the jump in positions would be unrealistic, and likely nausea inducing. For a more natural effect, we showed the user travelling between points over 1000 milliseconds. After trial and error, this was the right balance between showing clear movement between points, but not frustrating the user with movement which was too slow.

Initially we used a C# coroutine to linearly interpolate the camera frames between the start and end vectors in the environment. However, the movement between positions did not feel entirely natural – one participant commented that movement *"felt robotic"* as the user did not slow down on drawing towards their final position. Consequently we adjusted our C# routine to yield new camera positions between start and end vectors based on sinusoidal interpolation: $\sin(0.5 \times \pi \times t)$.

This is best visualised through Figure 4.15. With linear interpolation, at $t = 0.5$ the user will be halfway between start and end positions. In comparison, with sinusoidal interpolation they will have covered 70% of the path. Crucially, the sinusoidal function contains a large amount of deceleration towards the end of the movement. This alleviated the robotic nature of the original movement path, and resulted in a pleasant *'ease out'* effect. An excerpt of the code used to enumerate camera positions based on sinusoidal interpolation is included in Appendix D.

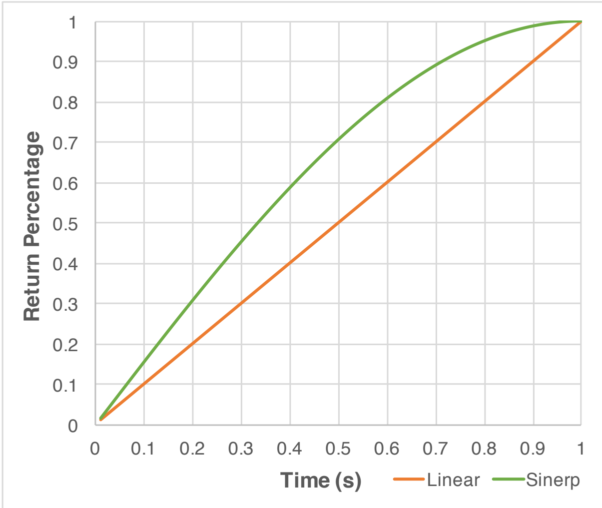


Figure 4.15: Linear and Sinerp Interpolation Comparison

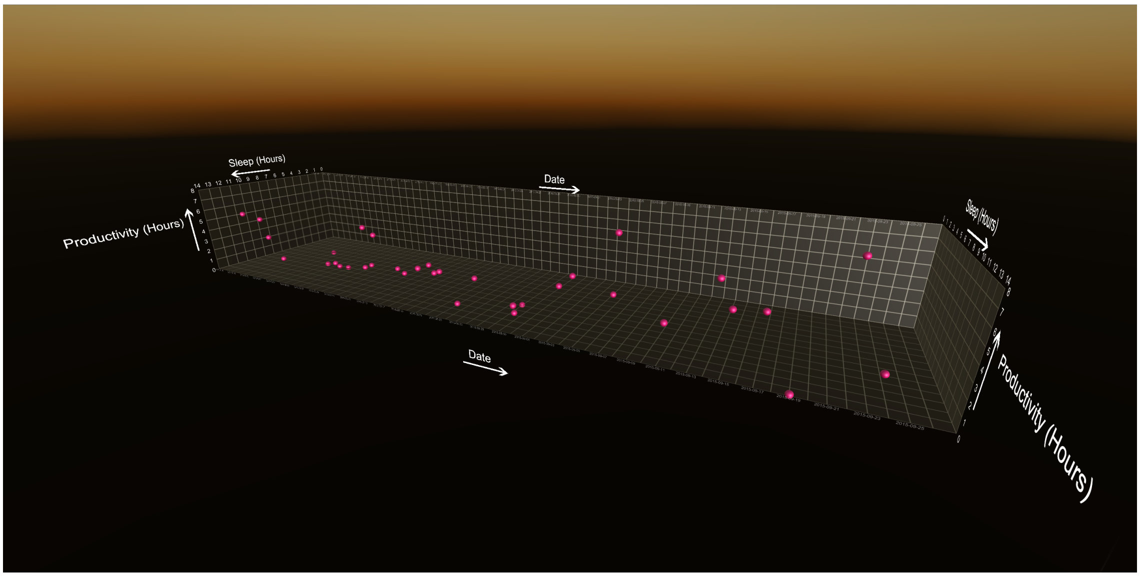


Figure 4.16: Final Prototype: Free movement in Be The Data

Free movement empowered users to view the graph in the Be The Data visualisation from new angles. With movement not constrained by the dataset, users had the ability to fly around the data, and make observations from a diverse number of perspectives. Figure 4.16 shows the view several participants took when making broader observations across the whole dataset. Prototype users also focused in on certain data points by using the controller to select and 'become' the data point. Typically this action resulted in local, deeper observations in our prototyping sessions. Images of this experience are included in

Chapter 5.

Free movement was also included in the Parallel Planes visualisation. Participants tended to move around less in this visualisation compared to Be The Data, but interacted more with the visualisation due to the brushing functionality which we developed.

Users had the ability to select multiple regions of the dataset to highlight certain data patterns against the entire dataset. Figure 4.17 shows an example where the user has selected all days where they had a mood rating of 5. Lines which intersect with the selected regions change colour to a vibrant pink, and non-intersecting lines are dimmed. This enables the user to identify if the brushed region – in this case days with an excellent mood rating – have any correlation with the other tracked dimensions over the dataset.

Figure 4.17: Final Prototype: Selection in Parallel Planes

In Chapter 2 we found that interaction is one of the core principles which makes VR an active learning experience, rather than a passive one. Through the movement and selection techniques discussed in this section, we have evolved the visualisations through prototyping towards "*natural, human-like interaction*" (Van Dam et al., 2000, p.41). The aim of this project is to see whether this immersive, interactive approach to visualising personal datasets will encourage new and different forms of user insight. The final section of this chapter considers additional design changes we made to our final prototype, in preparation for our user study in Chapter 6.

4.2.4 Further Design Considerations

The final section of this chapter is brief, and focuses on realism in VR and the missing data problem.

Realism

The Be The Data visualisation underwent many visual changes from the first to final prototype. Initially the spheres representing data points were coloured a shade between pink and blue, depending on their distance along the X axis. Our thoughts were that this may support users to perceive distance when looking along the X axis. However, prototyping feedback suggested that this caused confusion – some users believed that the colour of the sphere represented another data dimension altogether. Consequently we switched to a consistent vibrant pink for each data point. This complemented our next visual change which was to switch from lighting portraying midday to sunset in the environment. The pink colour popped well against the darker sunset background, making it clear for participants to delineate the data in the scene. This change was spurred on by a prototyping user mistaking the light source in the scene – the midday sun – as a data point.

These changes are relevant because of the context in which they happened. The Parallel Planes visualisation is not based in an environment with a form of natural light source. Furthermore, it does not contain a horizon, which acts as a constant point of reference in the Be The Data visualisation. The lighting and choice of environment affect the realism of this scene – the Be The Data visualisation appears to take place in some form of world, whereas there is no distinguishable world in the Parallel Planes visualisation. Through our experimental design we will measure the effect, if any, that realism may have had on data visualisation exploration for users.

Missing Data

The second design consideration to bring attention to is that of missing data. This is where a tracked user activity, such as the amount of sleep, is not recorded for a day or more. This has implications on the design of the visualisation, and how best to communicate unrecorded data to users. Originally in our first prototype of the Be The Data visualisation we used a non-sequential X axis recording date, where successive dates were not necessarily next to each other as empty dates were simply discarded. Quickly we realised that this was not easy to interpret, and in fact, the absence of a data point can often be an interesting observation in itself. In our final prototype we rendered an empty space where the data point for an empty date woudld have appeared.

Missing data was a more complex challenge for the Parallel Planes visualisation, and a challenge which we have not fully answered within the scope of this project. A drawback of the Parallel Planes visualisation is that as the number of dimensions increases, the greater the possibility that missing data will occur. This has negative consequences for the visualisation because if just one dimension is missing, the entire line for that day is affected – where does the line intersect with the plane for the missing dimension? A complex solution may be to draw a dotted line between non-affected dimensions such that the line can still travel across empty dimensions from the first plane to the last plane. However, this potentially adds visual complexity and, in the context of this project, was not feasible due to the additional development work. This solution was therefore not included in the scope of this project.

Instead, where data points were empty, we used the median value of the dimension up until that day. For instance, if sleep was not recorded on the 21st day of self-tracking, we used the median sleep time from the first 20 days to represent the 21st day. The median value was chosen over the mean, in the case that the dimension was not normally distributed and to negate the effects of outliers. Furthermore, we only used the median where genuine data was missing. For example, 0 music tracks played is not an instance of missing data – it could well be an interesting data observation. Ultimately the decision to use the median was a pragmatic one, made for the benefits of the user study in 6. We will be measuring how the visualisation affects insight generation on participants using a personal dataset which is not their own. Accuracy of the dataset would be more significant in a future system where users are exploring their own personal dataset.

4.3 Chapter Summary

In this chapter we have examined the design decisions made during the prototype phase. These prototypes build on the initial requirements list created during the previous chapter, and these requirements were refined through an iterative prototyping process described in this chapter.

We began by exploring dataset preprocessing, before discussing core design decisions with regards to axes and label representation, movement and selection, and empty data rendering within the visualisations.

Numerous figures were included throughout this chapter to elucidate these design decisions. The next chapter is almost entirely graphical and adds several more figures to demonstrate the working system which we will use during our empirical evaluation in Chapter 6.

Note: A code excerpt is included in Appendix D. All code is included in the zipped folder submitted through Moodle.

Chapter 5

Implementation Results

The purpose of this chapter is to visually demonstrate the design and functionality of our final prototypes. Screenshots of the visualisations are included to exhibit the systems which users will be interacting with in Chapter 6. The visualisations have been developed from the initial requirements list elicited in Chapter 3, and evolved through the prototyping process described in Chapter 4. Supplementing the figures throughout the chapter are general remarks on the fashion in which participants interacted with both visualisations during prototyping sessions. These comments provide the foundation for the following chapter in which we evaluate our visualisations systematically and more formally using an insight-based methodology.

5.1 Be The Data Visualisation

5.1.1 Starting View and Additional Perspectives

Figure 5.1 shows the opening view which users are introduced to when the Be The Data visualisation is started. The user is positioned outside the dataset, and half way down the X axis looking into the graph. The entire graph is not included in the user's field of view initially. This encourages them to look to either side to perceive the whole dataset.

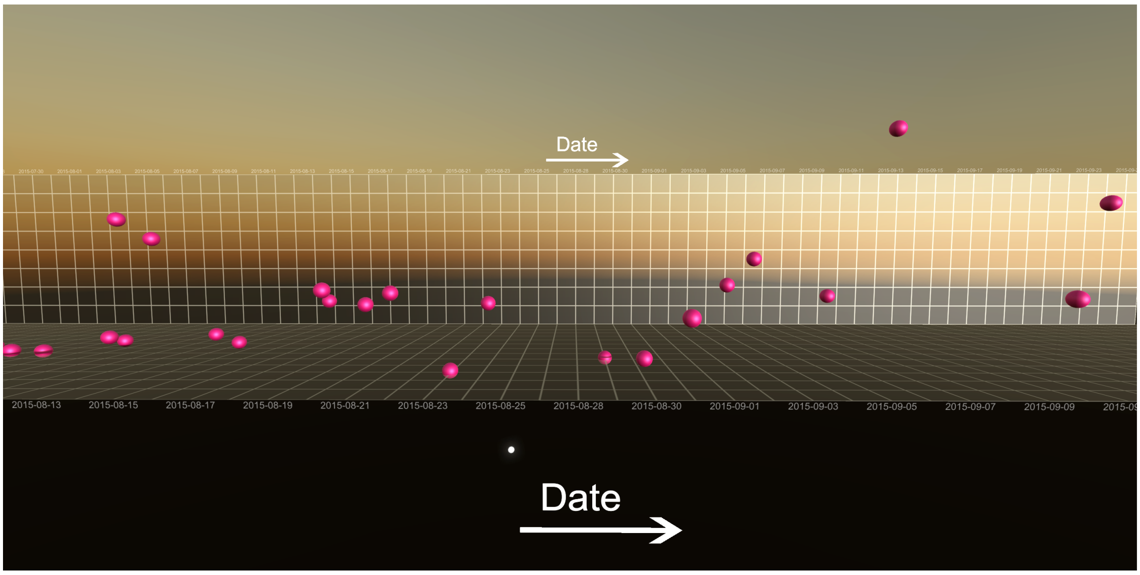
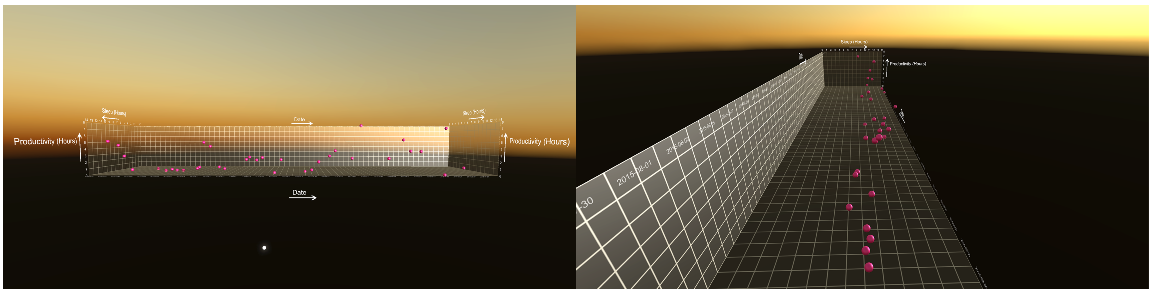


Figure 5.1: Be The Data: Starting Perspective

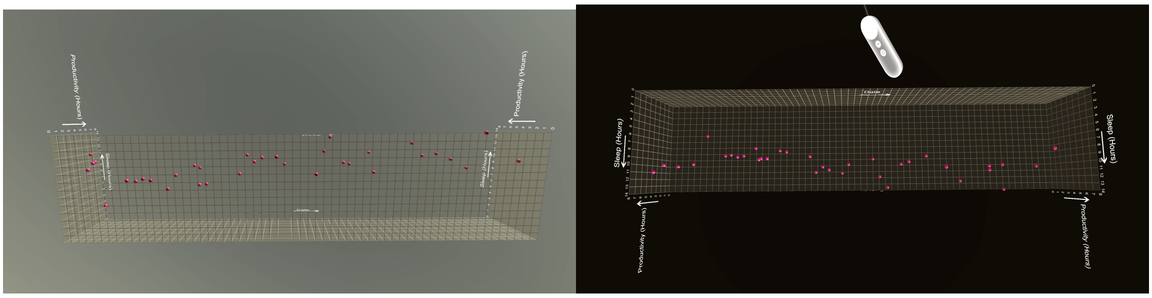


(a) Fully Zoomed Out

(b) Looking Down The Graph

Figure 5.2: Be The Data: Overview Positions

Typically in our prototyping sessions, users would now take one of two paths through the system. Each path involves making use of the swiping mechanism discussed in Chapter 4. Figure 5.2 (a) shows the first of these paths, where the user fully zooms out, allowing the entire graph to be observed in their field of view. The second path in Figure 5.2 (b) involves the user moving directly inside the graph and searching around in the visualisation tighter to the data points.



(a) From Below, Looking Up (b) From Above, Looking Down

Figure 5.3: Be The Data: Further Perspectives

Less common perspectives include moving directly above or below the graph and looking back at it. Prototyping participants almost exclusively moved to these perspectives towards the end of the prototyping sessions, if at all. Several participants realised that this in fact resulted in a 2D projection of the dataset, such as in Figure 5.3.

5.1.2 Looking Down the X Axis

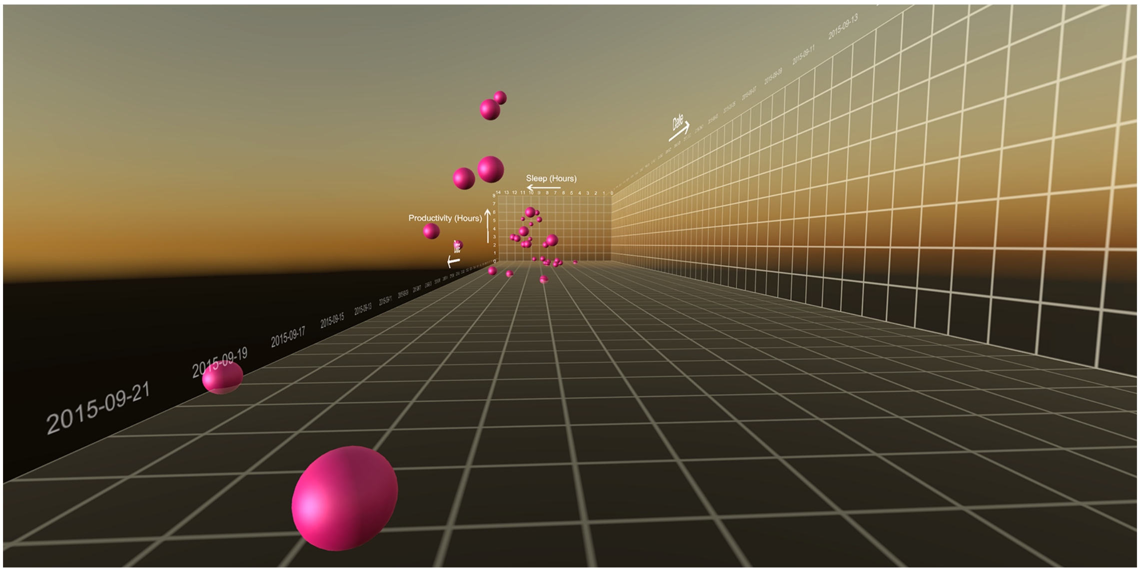


Figure 5.4: Be The Data: Looking Towards $X = min(X)$

A common task which prototyping users completed was to look down both directions in the X axis, towards the minimum and maximum X values. Introducing a plane at both ends of the X axis ensured that users were able to make 2D projections of the data against these planes. For example, in Figure 5.5 the user is able to see how *Productivity* (in the

Y axis) is related to *Sleep* (in the Z axis). By looking down the X axis which is plotting the *Date*, the user is able to see how *Productivity* and *Sleep* are correlated over time – represented by the distance of the X axis, or the depth of their forward-facing direction. Prototyping participants often completed this process by moving to one end of the graph and then moving forwards through the data. This enabled them to fly through the dataset, and observe how the 2D projection changed over time.

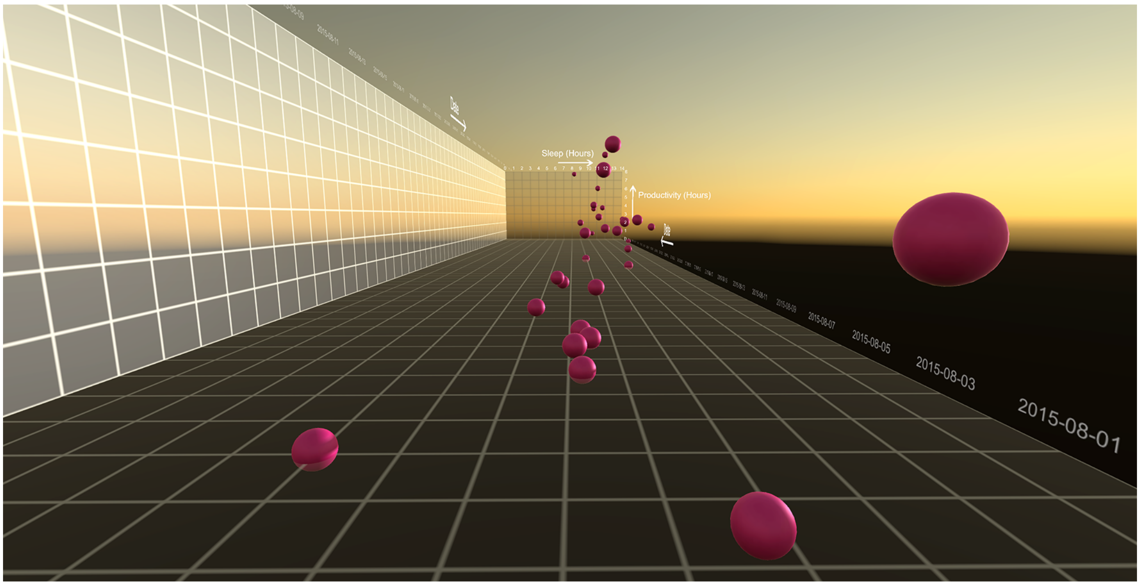


Figure 5.5: Be The Data: Looking Towards $X = max(X)$

5.1.3 Becoming A Data Point

In Chapter 2 we reviewed the Be The Data technique by Chen et al., in which students became data points in a physical environment. We identified requirements in Chapter 3 for users to become the data point in a virtual environment, and in turn that became the central concept of this visualisation. Users have the ability to point at data points with the Daydream controller (Figure 5.6) which produces an overlay about the selected data point. Users can then click on the data point to move towards it (Figure 5.7) and finally 'become' the data point (Figure 5.8). Users can use this mechanism to immerse themselves inside the data and travel between data points. Figure 5.8 shows some of the perspectives which becoming the data point can lead users to discover.

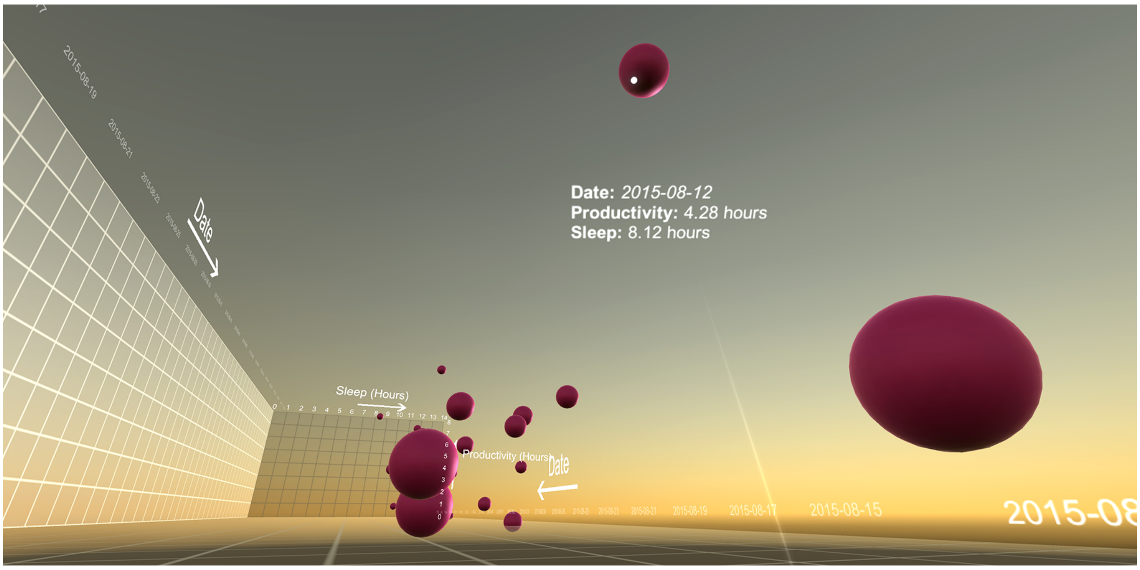


Figure 5.6: Be The Data: Data Point Selection

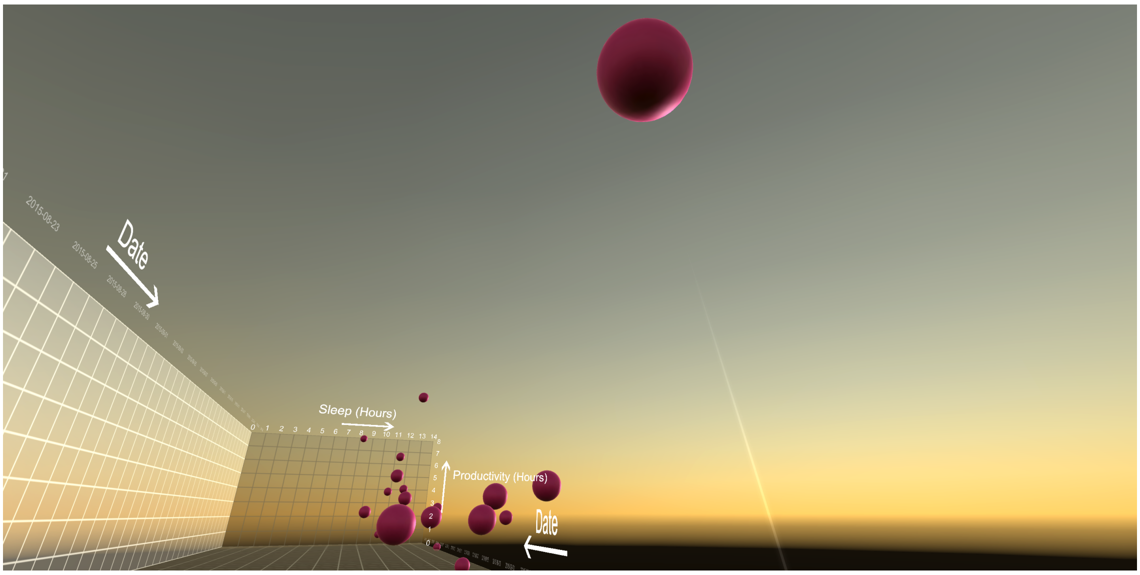
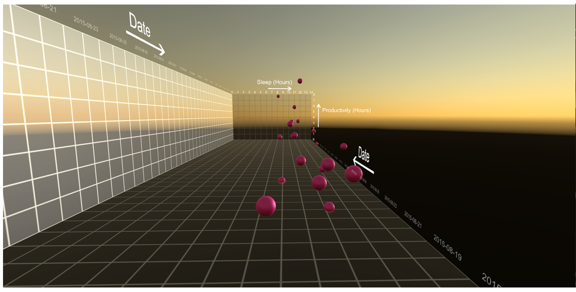
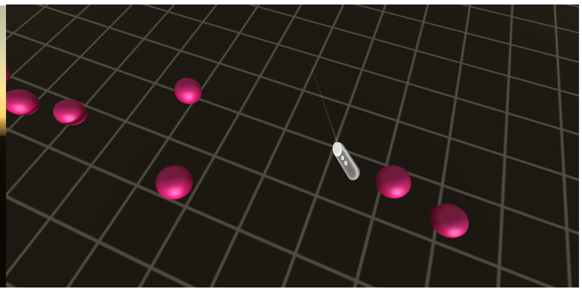


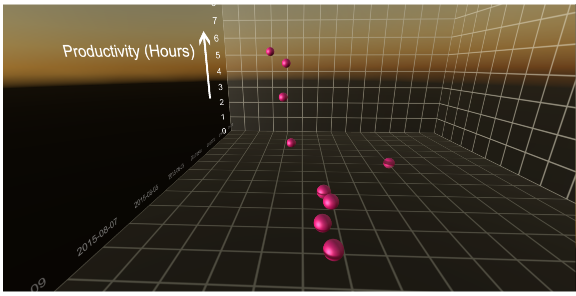
Figure 5.7: Be The Data: Travelling Towards The Data Point



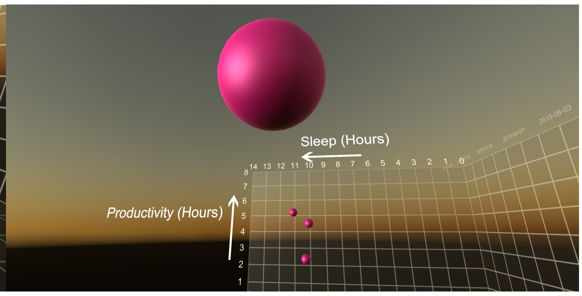
(a) Perspective 1



(b) Perspective 2



(c) Perspective 3



(d) Perspective 4

Figure 5.8: Be The Data: Looking Around From The Perspective Of The Data Point

5.2 Parallel Planes Visualisation

5.2.1 Starting View and Further Perspectives

The Parallel Planes visualisation begins with the entire dataset in view (Figure 5.9). Users are able to freely move around in the environment and view the planes from new angles. In our prototyping sessions we found that users typically did not move around a great deal, particularly in comparison to the Be The Data visualisation. A few users ventured around to directly face the planes (Figure 5.10(b)) and observed the dataset from this viewpoint. However, perspectives similar to Figure 5.10(c) and 5.10(d) were rarely used.

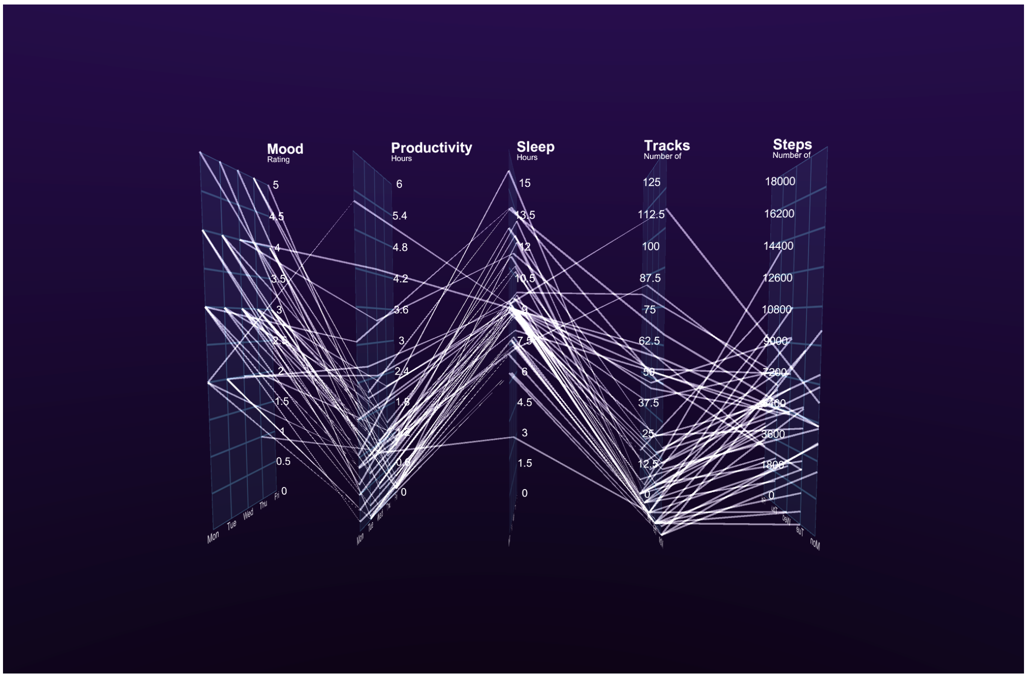
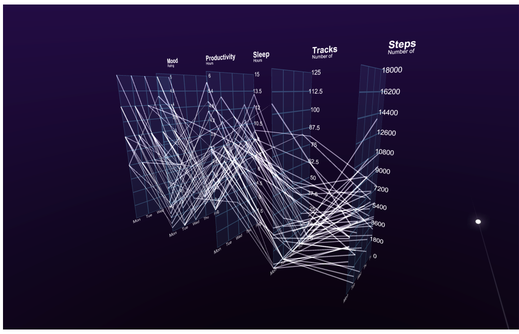
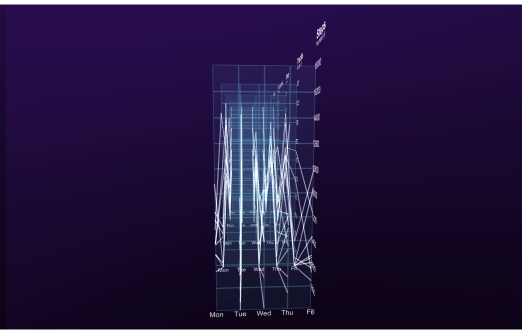


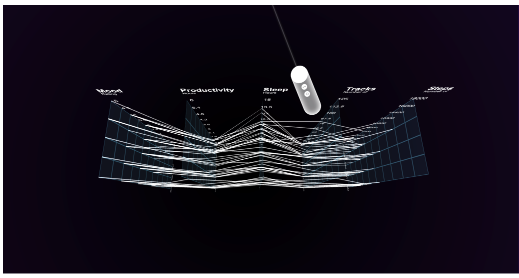
Figure 5.9: Parallel Planes: Starting Perspective



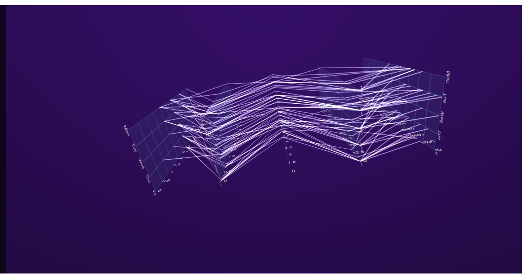
(a) Angled Perspective



(b) Side-On View



(c) From Above, Looking Down

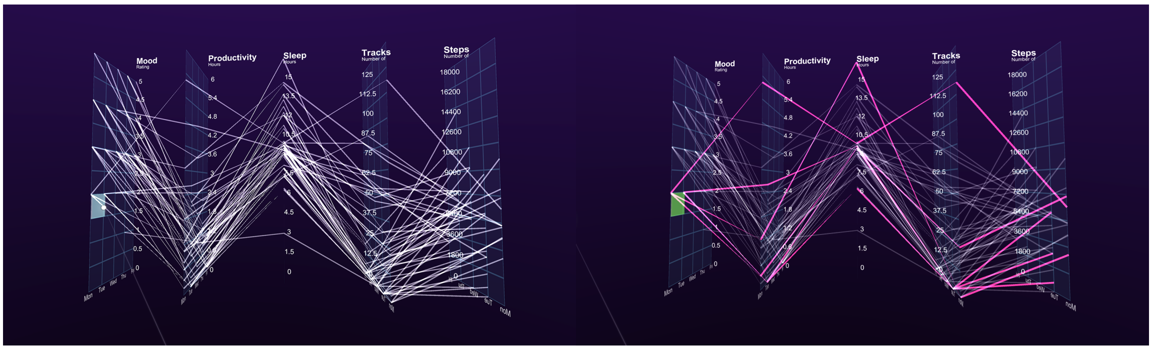


(d) From Below, Looking Up

Figure 5.10: Parallel Planes: Different Perspectives

5.2.2 Brushing via Selection

While the Be The Data visualisation encouraged movement, the Parallel Planes visualisation encouraged selection. Users were much more active in directly interacting with this visualisation in our prototyping sessions. We allowed users to select specific lines in the visualisation to see the distribution of the line over all dimensions. This used a similar brushing technique to the one created for the Parallel Planes visualisation by Brunhart-Lupo et al. (2016). However, our technique differed in that it allowed users to select and deselect squares on planes, rather than using a brush-like tool to paint regions over an entire plane. Figure 5.11(a) shows the result of hovering the Daydream controller over a square in our implementation, and in Figure 5.11(b) lines intercepting with this region have been highlighted after a controller click. This feature was compelling in our prototyping sessions – participants frequently used the technique to draw out subsets of the data and connect data observations together.



(a) Hover (b) Selection

Figure 5.11: Parallel Planes: Hover and Selection

We allowed users to make multiple selections, both on the same plane (Figure 5.12) and across multiple planes (Figure 5.13). Prototyping participants regularly made use of this ability to detect trends across all dimensions. For instance, the user could select regions relating to days of the week and see how the distribution across the dataset is affected. Similarly, the user could select across a data dimension. For instance, in Figure 5.12 badly rated days with a *Mood* rating of 2 are highlighted, and the user can begin to see how a bad day might affect other dimensions of data such as *Sleep* and *Productivity*.

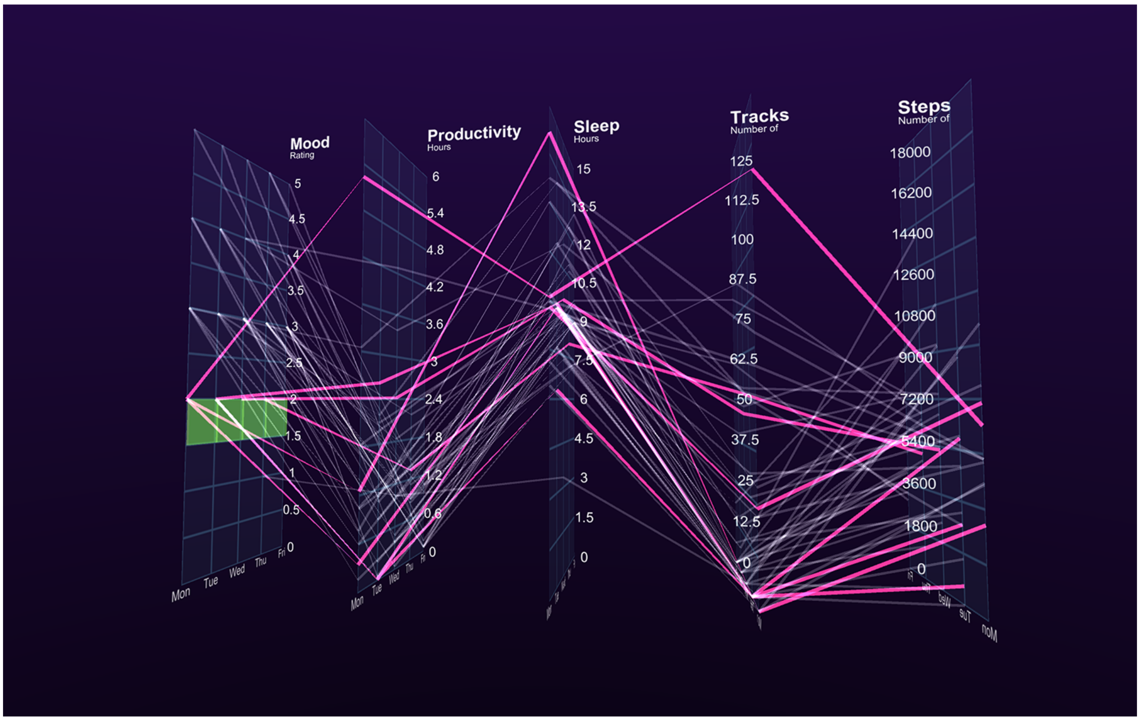


Figure 5.12: Parallel Planes: Multiple Selections

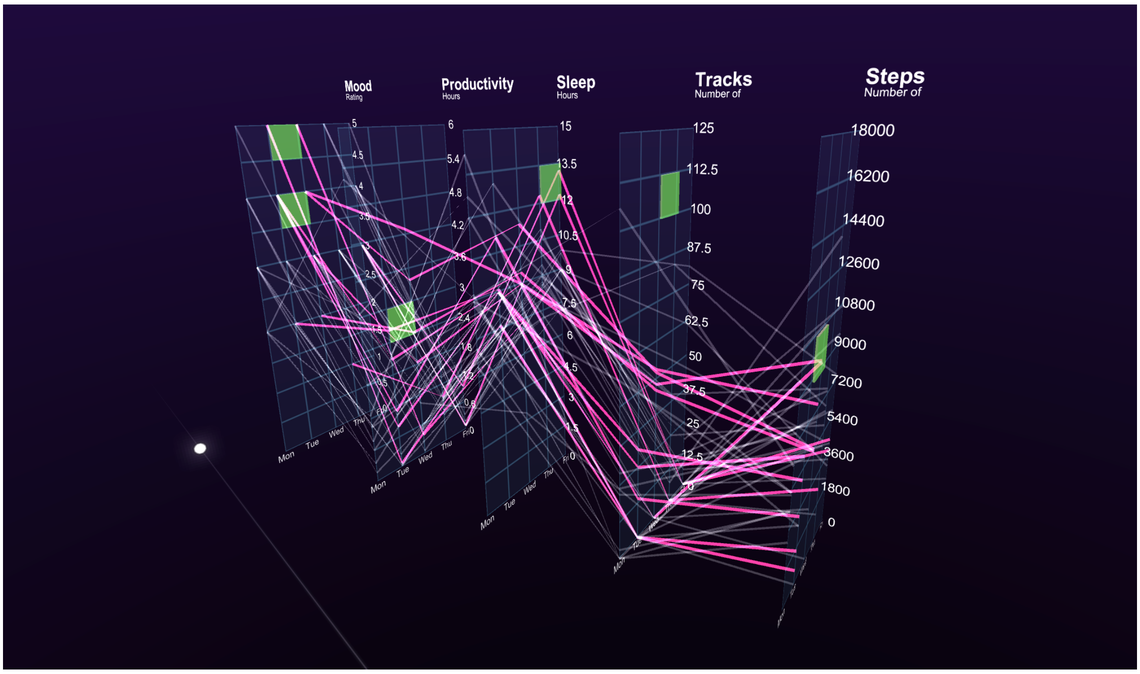
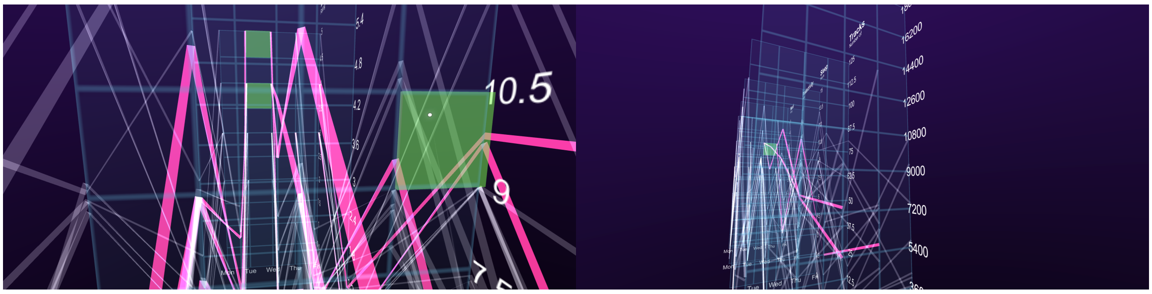


Figure 5.13: Parallel Planes: Multiple Selections Across Planes



(a) Inside The Visualisation (b) Side-On View

Figure 5.14: Parallel Planes: Different Selection Views

Finally, users were able to move around the visualisation while the selection was active. Several users used the side-on perspective at this point to confirm or reject ideas they had about the selection region. Several prototyping participants even travelled inside the planes (Figure 5.14(a)) and selected additional regions of interest. Nevertheless, observations about the dataset were generally made from a perspective in the visualisation where the entire dataset was in view (such as in Figure 5.12).

5.3 Chapter Summary

In this chapter we have given an outline of the typical approach in which prototyping users experienced our two visualisations. The Be The Data visualisation encouraged our prototyping participants to move around the graph for an overview of the dataset before selecting specific data points where they could immerse themselves in the data. Users traversed the dataset by selecting new data points, or by simply using the controller swiping mechanism to fly through the graph. In general, prototyping participants tended to take up positions in which they could observe the dataset in a 2D manner.

Comparatively, the Parallel Planes visualisation promoted the use of the selection functionality to pinpoint certain subsets of the dataset. Prototyping users moved around less in this environment, but spent more time directly interacting with the visualisation and seeking the discovery of correlations between new dimensions in the higher-dimension dataset.

With our prototyping process complete, and informal observations collected about the general behaviour of users interacting with our visualisations, we progressed to designing a study to thoroughly evaluate our implemented visualisations. In the following chapter we will describe our insight-based approach to evaluating these visualisations as a medium for personal data exploration.

Chapter 6

Experimental Design

This chapter will detail the approach we took to evaluate our visualisation in terms of three dependent variables: presence, task workload, and insight. First, we revisit the literature in order to compare different evaluation strategies, selecting questionnaires for presence and task workload, and Saraiya et al.'s insight-based approach for measuring visualisations.

We formally state research questions and hypotheses, and then review the variables which we measured and changed throughout our study. The design of the traditional 2D visualisations is discussed, and our experimental procedure is refined through a pilot study. Concluding this chapter is a summary of ethical and study materials, plus a description of the participant demographic.

6.1 Usability Studies

We begin by revisiting the literature to consider how our high-level dependent variables are traditionally evaluated. During the literature review we established that the principle purpose of visualisation is insight. Visualisation should generate insights for users working with existing hypotheses, and an effective visualisation will also support the identification of new hypotheses. As the following sections will discuss, an insight-based methodology was therefore required to evaluate the system successfully in terms of this criteria. Furthermore, we will discuss the approaches we took towards evaluating presence and workload, such that the study adopts several models of evaluation.

Presence

Presence – the sense of being in an environment – is very much associated with *sensory, cognitive and physical affordances*, terms developed by Hartson (2003). In the literature review, we argued that systems must maximise their influence along Sheridan (1992)'s principal determinants of presence. Sensory affordances along these three axes build towards

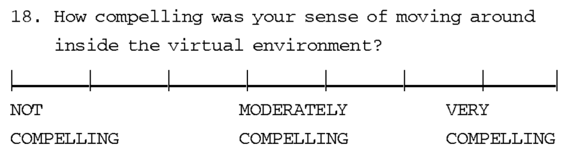


Figure 6.1: Example Presence Questionnaire Question (Witmer and Singer, 1998)

the notion of perfect presence. This enables users to see, feel, listen to and control various human senses. Sensing cognitive affordances is therefore crucial for participants to understand and reflect on their personal data, and sensing physical affordances is essential for participants to act upon this understanding (Hartson, 2003).

However, cognitive processes which sit between the immersive properties of a system and the users' perceived level of presence can become a moderating factor. Directly measuring the system would not yield an assessment of presence – presence is a psychological construct. Similarly, an objective measure of presence is difficult to take. The mental concept of perfect presence is not yet fully understood, and accordingly it is complex to quantify. *"Subjective report is the essential basic measurement"* (Sheridan, 1992, p.3) and so our assessment was completed through a questionnaire.

A users' subjective experience of presence has typically been evaluated through questionnaires in prior literature which assesses presence. Witmer and Singer (1998) constructed a questionnaire based upon previous literature, consisting of $N = 32$ questions (later reduced to $N = 19$ for reliability) using an adapted seven-point version of the semantic differential scale format. A cluster analysis was completed on the questionnaire data, labelling the factors of interaction and immersion necessary for participants to experience presence. A system resulting in 3 labels (Involvement/Control, Natural, Interface Quality) was conceived. A sample question from this example study is included in Figure 6.1.

Lessiter et al. (2001) later taxonomised existing studies on presence, finding virtual environments had multi-dimensional labels of presence discerned through factor analysis. Indeed, Schuemie et al. (2001, p.193) builds a comprehensive framework around existing studies of presence, concluding that measuring presence is *"almost exclusively"* completed through questionnaires.

After reviewing the presence questionnaires available to use, we chose the Igroup Presence Questionnaire (IPQ) for use in our evaluation. This is a 14-item questionnaire which fits closely with Sheridan's principal determinants of presence. The questionnaire was also chosen due to its length, which let us measure an appropriate amount of factors within the time available to us. A factor analysis of the questionnaire shows dimensions of spatial presence, involvement and experienced realism (Schubert et al., 2001), which we report our results in terms of.

Task Workload

A further comparison between 2D data visualisations and VR visualisations was made. This measured the participants' perceived workload of each system using the NASA Task Load Index questionnaire. More specifically, the *Raw TLX* was used in our study. This is the practice of dropping the second part of the questionnaire (Hart, 2006), which made it simpler to apply within the time constraints of this project. According to Hart (2006)'s meta-analysis, there is no definitive conclusion on the effect, if any, this has on the results of the questionnaire – multiple studies have found it to be less, more, and equally as sensitive as the full NASA-TLX. Accordingly we dropped the second part of the questionnaire, allowing a measure of task workload to be recorded within the time available.

The first part of the questionnaire asks participants to rate their answers to questions along six subscales: *Mental Demand, Physical Demand, Temporal Demand, Overall Performance, Frustration Level,* and *Effort.* The ratings are combined to give the overall Task Load Index. This is the participants subjective rating of their perceived workload when exploring the system, and will be used to give an indication of the effectiveness of the visualisation and its exploratory medium.

Insight

For the third evaluative approach, we used an insight-based methodology to quantitatively target participants interaction with the visualisations. While task performance is a frequently used measure for evaluating visualisation systems, it can be argued that predetermined tasks do not capture real-world data exploration scenarios which are often less guided and more exploratory. As a result, a task-based methodology was not used to evaluate this system. Instead, an insight-based methodology was be used. This enabled us to record and analyse insights during the participants open-ended exploration of the visualisation.

Saraiya et al. (2005) developed the insight-based evaluation methodology allowing visualisations to be measured beyond directed data analysis tasks. In particular, they defined an insight as an individual observation of the data visualisation and recorded insights (through a think-aloud protocol) as the participant discovered them. The insights were then analysed in terms of characteristics, producing 8 quantifiable attributes for each insight. While the study took place in a biological context, the evaluation method was not domain specific and so was applicable to this study. A list of insight characteristics determined by Saraiya et al. (2005) is summarised below:

- **Observation:** The data finding made by the participant.
- **Time:** The time taken to reach this insight. *Note: We dropped this measurement during our evaluation due to time constraints.*
- **Domain Value:** The significance or value of the insight. Saraiya et al. (2005)

encoded domain values on a scale of 1 to 5. Trivial observations resulted in a small domain value, and insights reasoning about new or existing hypothesis resulted in a greater significance.

- **Hypotheses:** Insights may result in participants identifying new hypotheses.
- **Directed vs Unexpected:** A directed insight answers a specific question that participants may have been asked to think about before starting the data exploration. Unexpected insights are observations discovered without being specifically searched for.
- **Correctness:** The correctness of the participants observation. Determined by domain experts.
- **Breadth vs Depth:** A breadth observation provides an overview of the entire dataset. A depth observation is detailed, and concentrates on a small number of data points. Determined by domain experts.
- **Category:** All insights should be analysed after collection and grouped appropriately. During an initial prototype phase, Saraiya et al. (2005) determined 4 categories for their domain: *Overview, Patterns, Groups* and *Details.*

Saraiya et al. (2005)'s insight-based approach was adopted in our study. Significantly, this enabled us to count distinct insights from participants, categorise them and analyse them quantitatively. This evaluative approach later enables us to reason about the effectiveness of the visualisations, and the immersive tools used for interaction with the visualisations. Rich, and more numerous, perceptions suggest an effective visualisation from which users can gain new and meaningful insights. It did not moderate our evaluation to a closed set of questions testing specific functionality of the system. Rather, it allowed participants to open-endedly explore the visualisations, generate insights, and create new hypotheses to investigate.

Our approach to the evaluation of our prototypes was therefore three-fold. Post-experiment questionnaires were composed of the IPQ questionnaire measuring presence, and a raw version of the NASA-TLX questionnaire measuring task workload. We evaluated the insights and hypotheses generated by participants through an insight-based approach. Specifically, this resulted in a quantitative measure of insight, instead of a measure of task performance. This evaluation therefore used a mixture of strategies for evaluation, measuring the extent to which the user understood and engaged with the personal dataset in the VR environment.

The following section summarises the progress of the project up until this point, and formalises the aims and research questions which we took into our evaluation.

6.2 Aims and Research Questions

The first part of this study set out to understand the motivations and challenges of personal data collection. In Chapter 2 we found that motivations for collecting personal data were directed by different types of goals, specific to many forms of data. Active user engagement with collected data was critical for identifying correlations in the data, and reflection over this data supports goal-setting.

However, one of the barriers to data reflection was that users did not find the level of information traditional data visualisations provided to be enough. Research has therefore looked towards including further data in the reflection process, such that users can obtain more appealing inferences. Nevertheless, the introduction of extra data was found to be overwhelming and complex to interpret in traditional, dashboard-based aggregation platforms.

Consequently we explored distinct visualisation methods and identified the two visualisation techniques which we developed in our final prototype systems: Be The Data, and Parallel Planes. The Be The Data technique engages the user by immersing them as a data point as part of a wider dataset – in this sense, the user 'becomes' the data point. Our second prototype system, Parallel Planes, transforms a highly-dimensional dataset into a lower-dimensional space in which the user can select subsets of the dataset. These visualisations were chosen to target active, user engagement with the dataset and to explore whether users could interpret a high-dimension dataset using the non-traditional reflection medium of VR.

With insight identified as the principal purpose of visualisation, it was clear that the insight-based methodology discussed in section 6.1 was required for the evaluation of these prototypes. To provide a clear direction for this methodology we revisited our initial aim and objectives, and refined them in terms of more specific questions which we were interested in answering.

The high level project aim of '*Is Virtual Reality a suitable medium for exploring personal datasets in?*' was decomposed into a series of three research questions. Thus, the purpose of the evaluation was to explore these three research questions through our user study and discussion of results.

- **RQ1:** Is the perceived workload for data exploration in VR different to exploration using traditional paradigms?
- **RQ2:** Does the use of VR in either *Be The Data* or *Parallel Planes* visualisations affect the insight generation process?
- **RQ3:** Does the *Parallel Planes* visualisation support the interpretation of a high-dimension dataset for non-expert users using VR?

To address these research questions, this chapter details how we evaluated our prototypes using the evaluation strategies listed in section 6.1. A wider discussion justifying the appropriateness of these methods took place during this section.

Summary of Experimental Process

In summary, our evaluation used the following methods:

- Saraiya et al. (2005)'s Insight-Based Methodology
- The Raw NASA Task Load Index Questionnaire
- The Igroup Presence Questionnaire (IPQ)

We did not have any specific tasks which we asked users to complete during the experiment. It was therefore an exploratory approach towards evaluation, well suited to Saraiya et al.'s insight-based methodology. This process enabled us to capture and analyse insights from users as they explored our prototype visualisations in an open-ended manner.

Two questionnaires were given to participants when they felt that they would not be able to gain any additional insight from the visualisations. The NASA TLX questionnaire gave an indication of the user's perceived workload for the open-ended data exploration. The resulting workload value was further divided into six separate subscales, which expressed the factors leading to the final workload measure more precisely.

Finally, the IPQ questionnaire was given to participants who report along three subscales of spatial presence, involvement, and experienced realism. There is a high factor loading on spatial presence which we used as a very general indicator of the success of the VR visualisations. This questionnaire was not given to participants in the control groups who used equivalent, traditional 2D visualisations. Instead, we used the findings from the VR experiments to contextualise and report on the results from both the insight-based methodology and raw NASA TLX questionnaire.

With research questions stated explicitly, and the aims and methodology of the study outlined, the following sections in this chapter presents further detail on how we evaluated our prototypes.

6.3 Hypotheses

The following hypotheses were derived and used in our experiment. These hypotheses were shaped by our choice of evaluation method and guided by our research questions.

6.3.1 Experimental Hypotheses

- **EH1:** There is a significant difference between the Task Load Index for participants in the VR visualisations and equivalent 2D visualisations.
- **EH2:** There is a significant difference between the number of insights generated by participants in the VR visualisations and equivalent 2D visualisations.
- **EH3:** There is a significant difference between the number of correct insights generated by participants in the VR visualisations and equivalent 2D visualisations.
- **EH4:** There is a significant difference between the number of unexpected insights generated by participants in the Be The Data visualisation and Parallel Planes visualisation.
- **EH5:** There is a significant difference between the number of hypotheses generated by participants in the Be The Data visualisation and Parallel Planes visualisation.

6.3.2 Null Hypotheses

The experimental hypotheses imply the following null hypotheses:

- **NH1:** There is no significant difference between the Task Load Index for participants in the VR visualisations and equivalent 2D visualisations.
- **NH2:** There is no significant difference between the number of insights generated by participants in the VR visualisations and equivalent 2D visualisations.
- **NH3:** There is no significant difference between the number of correct insights generated by participants in the VR visualisations and equivalent 2D visualisations.
- **NH4:** There is no significant difference between the number of unexpected insights generated by participants in the Be The Data visualisation and Parallel Planes visualisation.
- **NH5:** There is no significant difference between the number of hypotheses generated by participants in the Be The Data visualisation and Parallel Planes visualisation.

All of our hypotheses are two-tailed. Based on theoretical work considered in Chapter 2, which included a discussion of the immersive characteristics in VR and the multidimensionality of the Parallel Planes visualisation, we predicted that there would be an effect in all of our hypotheses. However, given that this study was exploratory, and with little existing research to based our hypotheses upon, we were not confident enough to specify direction through one-tailed hypotheses.

Accordingly, EH1 was chosen without a direction, and aimed to answer RQ1 – determining whether the data exploration workload was different between exploration mediums. EH2 and EH3 were chosen specifically with RQ2 in mind. The number of insights and the number of correct insights are key indicators that the insight-generation process is different between exploration mediums. The results from EH4 and EH5 will also be used to provide additional context for EH2 and EH3.

EH4 and EH5 were chosen to answer RQ3. We constructed EH4 as we predicted that the multidimensionality afforded by the Parallel Planes visualisation would lead to participants observing correlations they may not have been looking for initially. Saraiya et al. (2005, p.3) suggests that hypotheses are *"are most critical because they suggest an in-depth data understanding"*, and furthermore that hypotheses lead analysts toward "*continuing the feedback loop of the experimental process*". For these reasons, EH4 and EH5 imply that the participant can indeed interpret the visualisation, and has a deep understanding of the dataset. Consequently, these experimental hypotheses were selected for RQ3.

Finally, we did not formulate any hypotheses relating to presence. Due to the study format, only participants who interacted with the visualisations in VR received the IPQ questionnaire. As we predicted that we would not have enough results to perform statistical tests for this dependent variable, we did not create any hypotheses and instead we report on these results.

6.4 Variables

This section outlines the main variables which we focused on during the study. These are the key variables which we defined, manipulated and measured.

6.4.1 Independent Variables

These are the variables which were changed over the course of the experiment. There are two independent variables, each with two levels.

- Exploration Medium (*Levels:* Traditional 2D, VR)
- Visualisation (*Levels:* Be The Data, Parallel Planes)

6.4.2 Dependent Variables

These are the variables which were measured after changing the independent variables defined above.

- Insight
- Task Workload
- Presence

These variables are given at a conceptual abstraction level. Below we operationalise these variables in terms of specific, measurable statements:

- **Insight**
  - Breadth vs Depth – whether an observation provides an overview of the entire dataset, or is concentrated on a small number of data points.
  - Directed vs Unexpected – the insight may answer a specific question identified before exploration, or it is an observation discovered without being specifically searched for.
  - Hypotheses – the insight may result in participants identifying new hypotheses.
  - Domain Value – the significance or value of the insight. After analysing the collected insights we established the following schema:
    * 1: *General Observation.* e.g Mood is correlated with sleep.
    * 2: *Weighted or Specific Observation.* e.g Mood is positively correlated with sleep in these two weeks.
    * 3: *Explained Observation.* e.g The lack of pattern in sleep suggests they have a consistent amount of sleep, which also suggests that they go to bed at a reasonable time.
    * 4: *Hypothesis.* e.g I wonder if you do more steps, are you happier? (Given when participants identify a new hypothesis, or when answering a previously identified hypothesis not included in their initial exploration questions)
  - Correctness – the correctness of the observation. We labelled insights as *Correct, Neutral,* or *Incorrect.* Certain observations were marked as neutral due to subjectivity, and also in cases where hypotheses were identified.
  - Category – insights are grouped according to their type. We reused the categories provided by Saraiya et al. (2005) as they fitted our collected insights appropriately. These are *Overview, Pattern, Group,* and *Detail.*
  - *Note:* Due to projects constraints we are not including timing measurements for each insight in our data analysis.
- **Task Workload**
  As per the NASA TLX questionnaire, we measured task workload along the following six subscales.
  - Mental Demand
  - Physical Demand
  - Temporal Demand
  - Performance
  - Effort
  - Frustration Level

- **Presence**
  The 14 items included in the IPQ questionnaire can be seen in Appendix I.2. Most generally these items measured presence along the following subscales:
  - Spatial Presence
  - Involvement
  - Experienced Realism

6.4.3 Control Variables

In order to preserve the validity of the experiment we identified potential confounding variables. Our aim was to reduce the effect of these confounding variables such that bias and variance on our dependent variables was either reduced or eliminated.

Dataset

The major control variable in this experiment was the dataset used in the visualisations. The choice of dataset could completely change the complexion of the visualisations, so we kept the dataset consistent between different mediums to avoid positive or negative confounding. The dataset for each visualisation was therefore the same between the exploratory mediums – a participant interacting with a VR visualisation saw the same dataset as a participant interacting with a traditional 2D medium. For reasons explained shortly, we used a training session before the main test. Consequently, we also ensured that the dataset experienced on both mediums was the same during the training sessions.

The datasets used between visualisations was different. With the research aims of this study, it would not have been appropriate to represent the high-dimensional dataset used in the Parallel Planes visualisation in the Be The Data visualisation. Therefore the dataset used in the Parallel Planes visualisation was different to the one represented in the Be The Data visualisation. Our research questions and hypotheses were driven with this in mind, and the results are analysed in the context of these different datasets. A discussion of the datasets used in the visualisations continues in section 6.5.

Environment and Equipment

The use of VR required additional attention in terms of the environment used by participants. VR does encourage participants to move within their environment, rather than remain stationary. Accordingly, our experiments took place where participants were not restricted by the amount of physical space they used. In addition, all participants were afforded a comparable amount of physical space for the VR experiments.

Similarly, the physical lighting conditions was equivalent between participants because bright lighting conditions could have caused light leakage inside the VR headset. We made sure that each participant received a comparable VR experience. Likewise, we kept environmental noise levels to a minimum – not only for the VR exploration medium, but also for the traditional 2D medium.

The traditional 2D visualisations were displayed on the same device for all participants. This was a laptop with a 13.3" screen (2560x1600 resolution). The use of a mobile device to display the control visualisations was considered, but ultimately rejected. The lack of screen estate would have likely had a considerable confounding effect on the results of this experiment. The VR visualisations were displayed using the same headset, controller and phone for all participants. This ensured that each participant interacted with visualisations displayed on the same phone with a comparable resolution (2560x1440) and consistent running performance.

Participants and Training Session

Participants did not require any technical background to be involved in this study. In our study format we included a training session to allow participants to get familiar with the exploration medium controls – whether that was through the VR headset and controller, or on the laptop for the traditional 2D visualisations. This was introduced to control the effects of participants who were already familiar with VR headsets, such that all participants were able to start from a more balanced position. Furthermore, this enabled participants to get accustomed to interpreting a dataset through the initial training visualisations.

We strove to consider a diverse and balanced mix of participants for our study. However, given the constraints of a university project, we collected several personal details about participants. This allowed us to place the study results in the context of the participants involved. Namely these personal details are: age, gender, and highest achieved mathematical-related grade.

Format and Ordering

In order to prevent ordering effects confounding the study, our post-insight experiment questionnaires were counterbalanced with each other. This was only applicable to the VR exploration medium, where participants were required to complete two questionnaires on task workload and presence. Half the number of participants for each visualisation in VR completed the TLX questionnaire followed by the Presence questionnaire, and vice versa.

6.5 Experimental Data

This section describes the process of selecting the data to display in our visualisations. We had four different experiments, each with a training phase. Therefore we required four different cuts of our database (D_1, D_2, T_1, T_2) which were assigned as follows:

- Be The Data Training: T_1 *(2D, VR)*
- Be The Data: D_1 *(2D, VR)*
- Parallel Planes Training: T_2 *(2D, VR)*
- Parallel Planes: D_2 *(2D, VR)*

Each visualisation had one dataset for its training session, and one dataset for the main experiment. This was consistent between the exploration mediums. As previously mentioned, there was a distinction between the datasets used in the Be The Data and Parallel Planes visualisations – T_2 and D_2 contained two and three additional dimensions respectively.

A limitation of this project was that participants were not interacting with their own self-tracked data. However, to some extent this made our experiment fairer – participants interacted with the same collection of datasets. Our strategy for selecting these datasets to display was not an entirely random process. The manual approach involved the following steps:

- Our database contained tracked data about 19 participants. We first filtered this down to participants who had tracked the most data attributes (e.g sleep, tracks, steps) for at least two months.
- We then filtered these potential candidates into those with the least empty data points, allowing us to use the most complete and accurate dataset as possible.
- We limited the length of the remaining candidate datasets to the first two months of self-tracked data. The initial two month period tended to be data-rich for all candidate datasets, with only sporadic tracking beyond approximately 60 days.

We were left with three candidates to use for our experimental data. For our visualisations we selected just one candidate, and used their data across the Be The Data visualisation and the Parallel Planes visualisation. Where data points were missing, in the Be The Data visualisation we spatially represented this. In contrast, for the Parallel Planes visualisation we calculated the median along this dimension up to the missing data point (see section 4.2.4 in Chapter 4). Finally, we curated a number of attributes which we felt participants may want to identify potential correlations between and chose these for our visualisations.

To conclude, the self-tracked dimensions which participants explored in each visualisation:

- T_1 - Steps, Tracks
- D_1 - Sleep, Productivity
- T_2 - Active Step Time, Distracting Time, Sleep Awakenings, Tracks
- D_2 - Mood, Productivity, Steps, Sleep, Tracks

6.6 Traditional 2D Visualisations

After selecting the datasets above to be used in our VR visualisations, we began development of the visualisation to be used by the control group. These 2D visualisations were evaluated alongside the VR visualisations.

6.6.1 Traditional 2D – Be The Data

Our final 2D Be The Data visualisation consists of three standard graphs on a single PDF. The training and experiment visualisations are included in Appendix E. We used the T_1 and D_1 datasets defined above, imported them in a CSV style into Microsoft Excel, and exported three different scatter graphs for each dataset. For instance, for the D_1 dataset this was *Daily Sleep, Daily Productivity, Sleep vs Productivity.* These graphs were exported and combined into a single vector PDF which was then displayed full screen on a laptop. The significance of having a vector PDF, rather than a rasterised version, was that participants were able to use the pinch-to-zoom gesture to magnify parts of the visualisation without a reduction in quality.

The 2D visualisation was tightly comparable with our Be The Data VR visualisation. The 2D scatter graphs were directly available for each dimension in the dataset, much like how the prototyping participants made 2D projections against each plane in the VR visualisation. Study participants were also able to pinch and zoom into each graph in a similar manner to the swiping functionality in the VR environment. To the best of our knowledge, the 2D and VR representations of the datasets were *informationally equivalent.* Our evaluation procedure explored whether they were *computationally equivalent* (Payne, 2003).

6.6.2 Traditional 2D – Parallel Planes

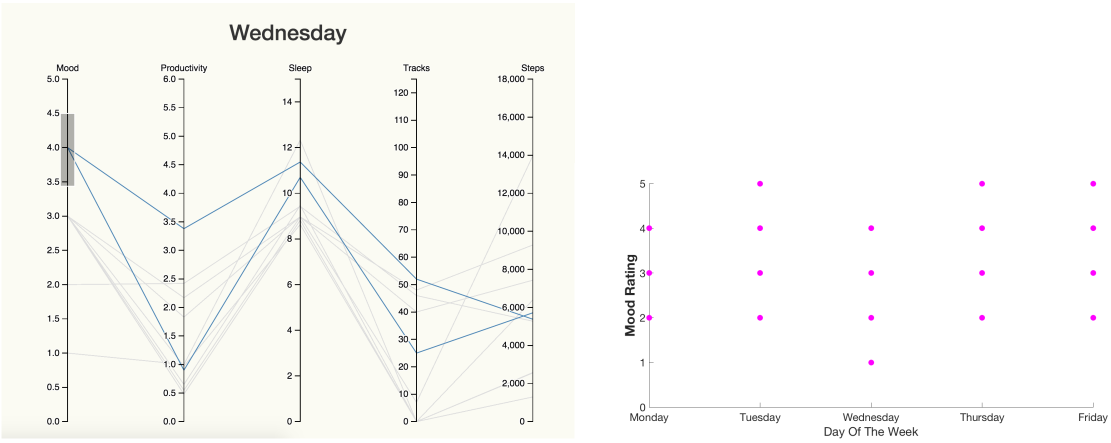
The 2D Parallel Planes visualisation contained five parallel coordinate visualisations representing each weekday, and side-on views for each dimension in T_2 and D_2. These were placed on a webpage implemented through a combination of HTML, CSS and the D3.js[1] Javascript library for the parallel coordinate visualisations. Each parallel coordinate visualisation represented the dataset dimensions for a particular weekday (e.g Sleep, Mood, Productivity on Monday). This meant that the coordinates were informationally equivalent – each parallel coordinate represented a unit of depth in the Z axis in VR. The parallel coordinates were displayed through an IFrame which enabled the participants to brush multiple visualisations separately.

For the side-on views, we wrote a Matlab script to plot a figure for each dimension in T_2 and D_2 against day of the week (Appendix F.1). The figures were then exported as images and included on the webpage. The justification for their inclusion was that in VR the

[1] https://d3js.org/

participant was able to move around in the environment and observe the Parallel Planes visualisation from the side. Each of our generated images was a flat representation of each plane in VR, generated to make the visualisations equivalent in information.

The Matlab and D3.js scripts read the T_2 and D_2 datasets in a CSV format. Similarly to the Parallel Planes visualisation where there is missing data, we calculated the median up to the current missing data point on the dimension. As a result, the datasets used between the 2D and VR Parallel Planes visualisations were identical. Finally the training and experiment webpages are shown to participants using a web browser in full screen mode – we used Python's SimpleHTTPServer[2] to serve these pages over a local web server.



(a) Example Parallel Coordinates (including brushing selection)

(b) Example Side-On View

Figure 6.2: Excerpts from the 2D Parallel Planes Experiment Webpage

The full training and experiment webpages are included in Appendix F. The study participant was able to pinch and zoom to specific visualisations on the webpage– for instance, the visualisations highlighted in Figure 6.2. Individual parallel coordinate visualisations could be brushed, much like how our Parallel Planes visualisation supported selecting and highlighting different subsets of the data. Functionally the only difference was that the brushing slider could be dragged up and down the axis on the 2D webpages, whereas in the Parallel Planes visualisation individual squares had to be clicked.

Appendix F shows that on the D_2 dataset, ten separate visualisations were required to represent the same amount of information in 2D as in our VR environment. Purely due to the structure of the data visualisations, it was not feasible to show all of these visualisations *above the fold* initially on the webpage. Consequently, there was a short scroll to view two additional visualisations, which are *below the fold*. Indeed, on a smartphone with even less screen estate, there would have been a great deal more scrolling. Our solution was the closest comparative solution which we were able to reach without significantly confounding

[2]https://docs.python.org/2/library/simplehttpserver.html

the experiment.

6.7 Experiment Setup

6.7.1 Format

We chose a randomised block design for this experiment, with the blocks determined by the participant's gender. This was not chosen with post-results data analysis in mind, rather to reduce variance across the results. In the scope of this project, we were not expecting to entirely remove gender as a potential extraneous variable due to the realistic number of participants we would be able to recruit. Nevertheless it did allow us to reduce data variance in a minor, peripheral way.

Each participant was assigned a unique identifier and explored one visualisation in this study. The following groups were identified to which participants were assigned:

- 2D Be The Data
- 2D Parallel Planes
- VR Be The Data – Presence then TLX questionnaire
- VR Be The Data – TLX then Presence questionnaire
- VR Parallel Planes – Presence then TLX questionnaire
- VR Parallel Planes – TLX then Presence questionnaire

These groups took into account the aforementioned counterbalancing of post-exploration questionnaires. Given that we required an equal number of participants to complete the experiments in 2D and in VR, this meant that the number of participants in the study could only be a multiple of 8. After being split into blocks, participants were randomly assigned to each group above and were only involved with a single condition. Repeated measures would not have been suitable for this study due to ordering effects. Most noticeably this would have been fatigue, and also participants becoming competent with the think-aloud insight-based approach.

6.7.2 Pilot Study

A short pilot study was conducted with 4 participants (2 Male, 2 Female). The pilot participants had not been part of the prototyping phase, so like future study participants, they were experiencing the visualisations for the first time. Each participant was assigned to an exploration medium (2D or VR) and a visualisation (Be The Data or Parallel Planes). The purpose of completing a pilot study was to validate our traditional 2D visualisations and final VR prototypes. Furthermore, to refine the study design and evaluation process such that we we would be able to effectively test our hypotheses.

After the execution of our pilot study, and subsequent data analysis, there were only minor changes which we made to our study structure and visualisations. These are summarised below:

- *Visual modification to the 2D Parallel Planes visualisation.* The pilot participant assigned to this experiment found it challenging to differentiate between the 5 different graphs for each weekday. Consequently we added unique, light background colours to each parallel coordinates visualisation on the webpage.
- *No audio recording of the training sessions.* Initially we recorded audio over the entire experiment, in line with the insight-based methodology outlined by Saraiya et al. (2005). However, transcribing each participant's observations into insights was a time-consuming process, particularly when both the training and main experiment stages were included. To reduce a potentially overwhelming amount of audio analysis, we limited the amount of audio we recorded to just the main experiment.
- *Verbally emphasise the questionnaires are about the main experiment.* After completing both the training and main experiments, when filling out the questionnaires two pilot participants clarified whether they were being asked about the training session and main experiment, or just the main experiment. In the main study we made it explicitly clear that the questionnaires were about the main experiment only.

The refined study method is laid out in the next section. This is the final procedure which we used during our main study.

6.7.3 Study Procedure

The following bullet point list was intended for use by the researcher in charge of the experiment. The study was split into four stages, and detailed below as to ensure the consistent running of the experiment between participants.

Participant Brief

- Show the participants the consent form. Explain the contents of the brief and answer and potential questions. Gather information on age, education, gender, and mathematical level.
- Explain each stage of the experiment. In particular, an informal explanation of the *think aloud* protocol.

Training Stage

- Explain the exploratory medium controls to the participant
  - *If VR:* Adjust the Daydream headset to the participant's head size and introduce them to the Daydream controller buttons.
  - *If 2D:* Explain the pinch to zoom gesture.
- Show participant the dimensions included in the training dataset and give a background description on each one (e.g unit, how it was measured).
  - Be The Data Training: Steps, Tracks
  - Parallel Planes Training: Active Step Time, Distracting Time, Sleep Awakenings, Tracks
- Get the participant to think of 3 questions they wish to ask of this dataset. State that these are only initial starting points for exploration.
- Introduce the exploratory medium, and state that there are no time limits for exploration. The participant should finish when they feel they will not be able to gain any additional insight.
- Periodically ask the participant for an estimate of how much of the total potential insight they feel they have obtained so far.
- Finish when the participant is not able to gain any additional insight. Ask the participant for a final total insight measurement.
- Make sure that the participant is comfortable with the controls of the exploratory medium.
- Briefly take away the exploratory medium. Switch from the training visualisation to the experiment visualisation.
- Talk through the training experience and make sure that they understand the interaction tools (e.g brushing in the Parallel Planes visualisation)

Main Stage

- Show participant the dimensions included in the experiment dataset and give a background description on each one. (e.g unit, how it was measured).
  - Be The Data: Sleep, Productivity
  - Parallel Planes: Mood, Productivity, Sleep, Tracks, Steps
- Get the participant to think of 3 questions they wish to ask of this dataset. State that these are only initial starting points for exploration.

- Start audio recording
- Introduce the exploratory medium, and state that there are no time limits for exploration. The participant should finish when they feel they will not be able to gain any additional insight.
- Periodically ask the participant for an estimate of how much of the total potential insight they feel they have obtained so far.
- Finish when the participant is not able to gain any additional insight. Ask the participant for a final total insight measurement, and any additional qualitative comments.
- Stop audio recording
- Take away the exploratory medium.
- *If VR:* In the required order for this participant, complete the Presence and NASA TLX Questionnaires.
- *If 2D:* Complete the NASA TLX Questionnaire only.

Participant Debrief

- Explain the objectives of the study and the questions we are asking. Answer any further questions which the participants may have
- Finally, thank them for their time.

(a) (b) (c)

Figure 6.3: Study participants in VR using the Daydream headset and controller

6.7.4 Ethics and Study Materials

The departmental 13-point ethics checklist was discussed with Dr Simon Jones, project supervisor, and is attached in Appendix G. This supported the creation of the consent forms, which also acted as briefing forms for participants. Two separate consent forms with a requirement for the participant's signature were created – one for traditional 2D visualisations, one for VR – and are included in Appendix H. The study procedure discussed in Section 6.7.3 was also printed to support the researcher running the experiment.

The multi-dimensional NASA TLX scale is included at Appendix I.1. This was provided to participants when they could not get any further insight from the visualisations, alongside a sheet containing definitions for each scale (Appendix I.1.1). Participants were asked to mark along each scale on the paper scale with the pen provided. The NASA TLX scale was obtained from the NASA website (Hart and Staveland, 1988*b*) and the explanatory definitions from the seminal paper (Hart and Staveland, 1988*a*).

The IPQ presence questionnaire was counterbalanced with the TLX questionnaire to avoid order effects confounding the experiment. The primary reference for the IPQ is (Witmer and Singer, 1998) and Schubert et al. (2001). The Igroup website[3] contains the 14 items which make up the questionnaire, including the Likert scale anchors. A Google Form[4] was created with the 14 items, and the participant filled in the questionnaire on a laptop after exploring their assigned visualisation. The presence questionnaire which participants completed is included at Appendix I.2.

6.8 Participant Breakdown

Our experiment used two blocks of 8 participants, resulting in a total of 16 participants. The sample was relatively young ($M = 21.6$, $SD = 0.60$) and was formed entirely of undergraduate students. 13 participants had achieved a Grade A$*$ or A at A Level Mathematics, or equivalent. The distribution of gender was 8 Female, 7 Male and 1 Prefer Not To Say. Accordingly, we split the participants into two equal blocks (Female, Male + Prefer Not To Say) and randomly assigned the participants in each block to an experimental condition.

The participants we recruited had not been part of our prototyping process, nor had they been part of the pilot study. Therefore, during the main study, participants would experience the visualisations for the first time.

In total, 8 presence questionnaires, 16 task workload forms, and 288 insights were collected from our participants. Each insight was analysed in terms of the 6 measurable attributes defined in Section 6.4.2.

[3] http://www.igroup.org/pq/ipq/download.php
[4] https://www.google.co.uk/forms/about/

6.9 Chapter Summary

During this chapter we revisited the literature in order to consider and design the three-fold approach used to evaluate our prototype visualisations. Specifically, we selected the IPQ questionnaire for measuring presence, the raw version of NASA's task workload questionnaire, and Saraiya et al.'s insight-based methodology for visualisation evaluation.

We then explicitly stated our research questions, and highlighted hypotheses, and independent, dependent and control variables. The datasets used in the study were detailed, and the designs of the traditional 2D visualisations were discussed. We concluded by reviewing the study procedure refined through our pilot study, explaining how participants were assigned to experiments using a randomised block design.

The next chapter contains the results from our study, alongside analysis, in order to fully evaluate the hypotheses we identified during this chapter.

Chapter 7

Results and Analysis

This chapter contains the results of our empirical evaluation. It is split into three sections corresponding to our three conceptual dependent variables: presence, task workload, and insight. The presence results are reported in order to contextualise the results on task workload and insight, which are statistically analysed in terms of our hypotheses. The chapter concludes by reviewing qualitative feedback from participants, organised by the visualisation and themes which appeared during the study.

Due to the size of the empirical evaluation, the discussion is split over this chapter and the following chapter. This chapter presents the results and analyses their implications at a low level of detail. The following chapter puts the results into the context of research question outcomes, and details the future research potential of the visualisations.

7.1 Presence Results

8 questionnaires were collected from participants who explored the *VR Be The Data* and *VR Parallel Planes* visualisations. The IPQ questionnaire has 3 subscales on its 14 questions which were identified through a factor analysis during its construction by Schubert et al. (2001). These are spatial presence, involvement and experienced realism. We will therefore report our results in terms of these factors.

We will not extensively analyse these results due to the limited sample size of 4 participants per experiment. The purpose of reporting these results is to capture an indication of the users' perceived presence and the potential influence this may have had on task workload and insight generation.

The IPQ questionnaire itself is available in Appendix I.2 and the raw results gathered are in Appendix K.1.

Spatial Presence

Q2-Q6 in Figure 7.1 indicate the range of responses for the spatial presence factor of the questionnaire. Q1 is included in this figure, but it is the exceptional item in this questionnaire. It is a general item added by the IPQ creators and it is not specific to one factor. It is loaded along all three subscales according to their factor analysis[1] – Q3 is reverse-coded on the Likert scale, which is why the response to this item stands apart from the other items.

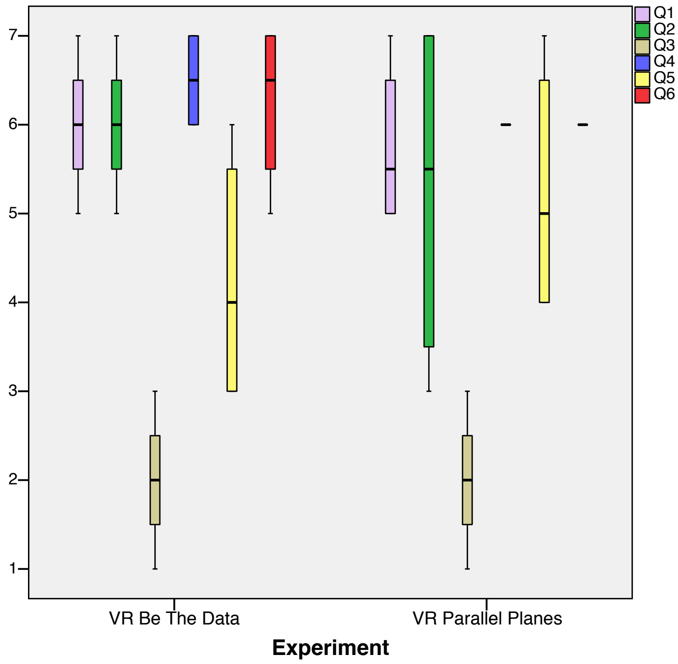


Figure 7.1: Spatial Presence Box Plot – Experiment Comparison

Q2-Q6 report similar results in terms of spatial presence between the two visualisations. On average, the Be The Data visualisation had a greater score on the Likert scale in every question, with the exception of Q5. This question was: *"I had a sense of acting in the virtual space, rather than operating something from outside"*. We anticipate that this was due to the proactive role which users took to brush the dataset in order to change its visual appearance.

The responses for Q2 had a larger range in the Parallel Planes visualisation. This item asked participants to indicate their agreement with the statement: *"Somehow I felt that the virtual world surrounded me"*. A reason for the variation in responses is perhaps due to the relative lack of movement in the Parallel Planes visualisation. We observed participants in

[1] http://www.igroup.org/pq/ipq/factor.php

the Parallel Planes visualisation remaining static, and exploring the dataset from a single vantage point throughout the experiment. Comparatively, participants actively moved around inside the Be The Data visualisation. In this experiment, participants looked around in all directions from several positions in the virtual world . In particular, when participants were inside the graph, the data was literally surrounding them in all directions, as opposed to a single perspective view for the Parallel Planes visualisation.

Involvement

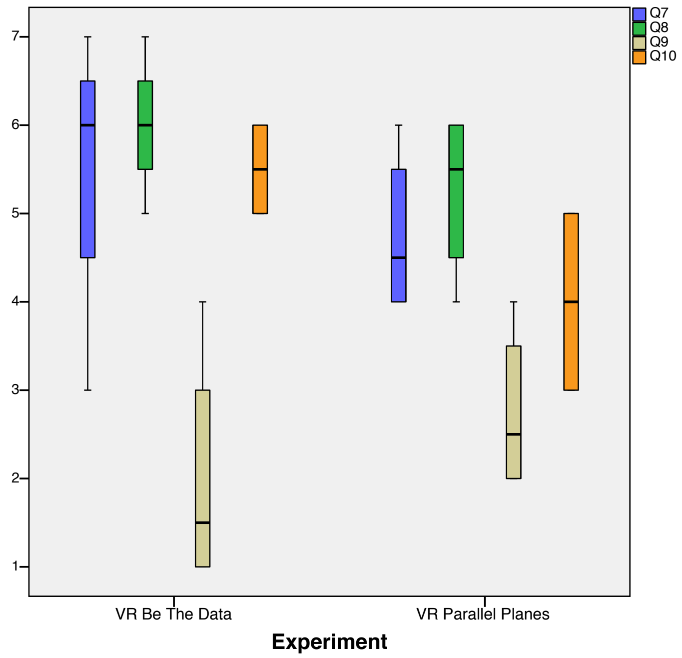


Figure 7.2: Involvement Box Plot – Experiment Comparison

The IPQ questionnaire makes the distinction between an attention component of presence, and a spatial cognitive component. The previous section reported on this latter component – this section will report on the attention component of Involvement given by Q7-Q10 (*Note:* Q9 is reverse-coded)

In general, participants reported that the Be The Data visualisation performed better on the Involvement factor. The questions asked participants about their attention (Q8: *"I was not aware of my real environment."*) and the extent to which the visualisation engaged them (Q10: *"I was completely captivated by the virtual world."*).

The cause of the difference between the two visualisations is not clear, although one may look towards the way in which the virtual scenes are presented in VR. The Be The Data visualisation is set in a world with a horizon, with a distinguishable ground and sky. Comparatively, the Parallel Planes visualisation does not contain any recognisable landscape. Instead, participants explore the dataset with a continuous dark purple background, and no reference to a horizon. The difference in results may have been affected by the participants ability to enter a state of psychological presence due to the appearance of the scenes in VR.

Realism

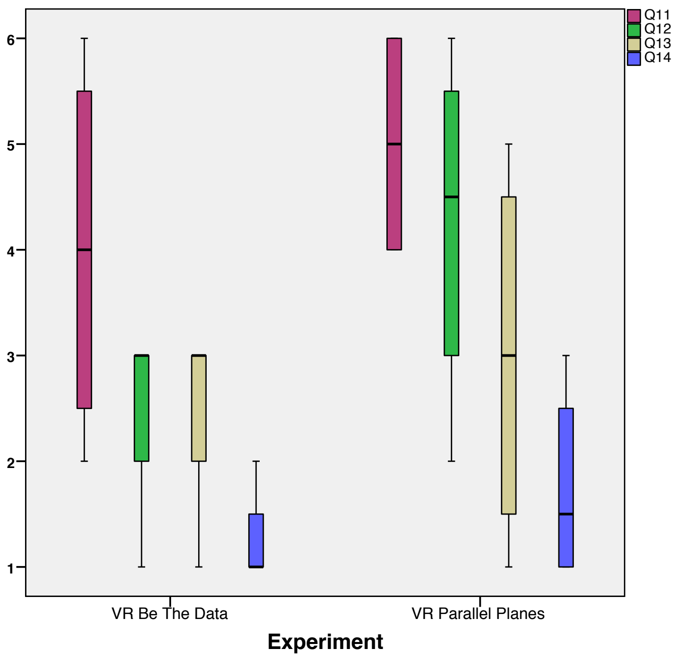


Figure 7.3: Realism Box Plot – Experiment Comparison

Participant's responses to questionnaire items had generally been above-average on the 7-point Likert scale for factors of spatial presence and involvement. However, for the final factor of realism, responses fell below average for both experiments.

Four questions relating to realism were asked to participants (Q11-Q14, with Q11 reverse-coded). Notably the Be The Data visualisation elicited a poor response in terms of items such as *"How real did the virtual world seem to you?"* and *"The virtual world seemed more realistic than the real world"*.

The visualisations presented to participants during the experiments did not rate well for realism. In Chapter 2, we discussed Witmer and Singer (1998)'s suggestion that consistency in information in virtual worlds and the real world leads to a gain in presence. One could hypothesise about the importance of realism in the data visualisation domain, particularly when designing such systems. As a design factor, should realism be placed at a similar level of importance as is done for VR systems across simulation and exposure therapy domains? With a lack of literature in our research area, it is not clear whether the poor realism results will affect our findings in the context of data visualisation analysis. Nevertheless, this limitation will be considered when analysing our other dependent variables of task workload and insight. Furthermore, this finding will be highlighted as part of our contribution in Chapter 8.

Further Discussion

Although the IPQ questionnaire results have shown that our visualisations are limited in terms of realism, our system performs well in the other factors of presence: spatial presence and involvement. In order to ensure that we had indeed measured the concepts of presence we wished to discuss, we calculated Cronbach's Alpha for each factor of the IPQ questionnaire. Q3, Q9 and Q11 were reverse scored to account for their initial scoring on the 7-point Likert scale.

The purpose of calculating these values was to verify that the presence questionnaire was a suitable evaluation method for our data visualisations. Cronbach's Alpha were calculated separately for dimensions of spatial presence (5 items; $\alpha = .69$), involvement (4 items; $\alpha = .73$), and realism (4 items; $\alpha = .74$). The combination of relatively strong a coefficients and comparable results on each subscale with example studies[2] indicated that the questionnaire items were consistent in measuring the concepts we were targeting.

We calculated the mean composite scores of presence of 60.3 ($SD = 3.0$) and 61 ($SD = 4.95$) for the Be The Data and Parallel Planes visualisations respectively. For these calculations we did not transform reverse-coded questions. This enabled us to compare our presence results against other VR systems in the literature (Morina et al., 2014; Zlomuzica et al., 2016; van Gelder et al., 2016). We found that our system performed favourably in terms of creating a sense of presence in comparison to existing systems who had also measured presence using the Igroup Presence Questionnaire. Nevertheless we were not able to make a true comparison in our research domain. An exhaustive search of the literature did not yield measures of presence for data visualisation in VR, let alone personal data visualisation.

Our final presence results show a close similarity for the participants' experienced presence in both the Be The Data and Parallel Planes visualisations. Participants in the Be The Data visualisation felt more involved in the virtual environment and paid less attention on average to the real world. Both visualisations performed poorly on the realism factor, but in spite of this had favourable overall presence measurements when compared with other VR systems in the literature.

[2] http://www.igroup.org/pq/ipq/factor.php

7.2 Task Workload Results

7.2.1 Testing EH1: Overall Task Workload

Our first hypothesis predicted that there would be a significant difference between the NASA Task Workload Index for participants in the VR visualisations and equivalent 2D visualisations. To test this, we combined participants across both visualisations into two groups of 8 participants each. This enabled us to evaluate the task workload significance across 2D and VR groups through a two-way ANOVA.

Three assumptions had to be assessed before we completed the two-way ANOVA. These were:

- **Assumption of Independence**: As defined by the randomised block design, participants were assigned to one experimental treatment only.
- **Assumption of Normality**: Data was formally tested for normality through the Shapiro-Wilk test on *2D Be The Data* (W = 0.870, dF = 4, p= 0.296), *2D Parallel Planes* (W = 0.861, dF = 4, p= 0.264), *VR Be The Data* (W = 0.912, dF = 4, p=0.492), and *VR Parallel Planes* (W = 0.985, dF = 4, p= 0.933). The Shapiro-Wilk significance values were all greater than 0.05 which allowed us to conclude that the data was normally distributed.
- **Homogeneity of Variance**: We used Levene's Test for Equality to check that the variances of our four groups were equal. Levene's Test was not significant: F(3,12) = 2.208, p = 0.140 at the 0.05 alpha level. Thus we concluded that the homogeneity of variance assumption was met.

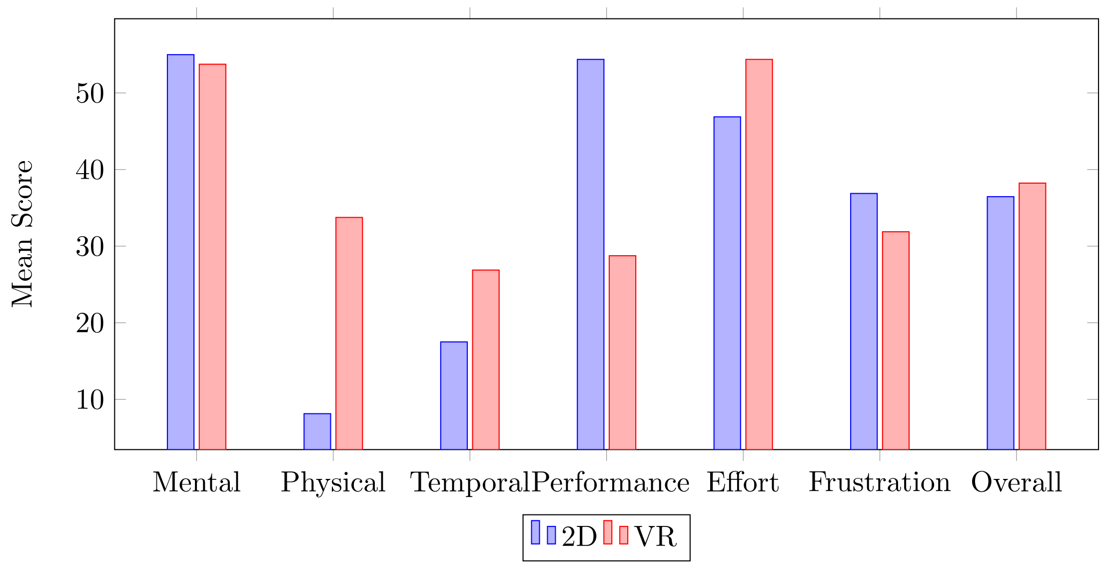
Since none of these assumptions were violated, we ran a two-way ANOVA to examine the effect of exploration medium and visualisation on task load index.

There was no statistically significant interaction between exploration medium and visualisation, $F(1, 12) = 0.03$, $p = 0.956$, nor was there a statistical difference on task load index due to the visualisation used in the experiment ($\alpha = 0.5$). Similarly, there was no statistically significant difference by exploration medium, $F(1, 12) = 0.104$, $p = 0.753$. Therefore the results for the 2D visualisations (M = 36.46, SD = 8.05) and the VR visualisations (M = 38.23, SD = 12.01) were not significant. Hence, we failed to reject NH1 given that $p > 0.05$ between 2D and VR mediums.

Since we failed to reject NH1, we were able to conclude that in our study that there was no significant difference between the task load index for participants in the VR visualisations and equivalent 2D visualisations. However, the use of a scalar measurement to represent a multidimensional measurement of task workload has been questioned by researchers previously (Hendy et al., 1993; Salvendy, 2012) and many studies analyse the task workload subscales individually, in addition to an overall measurement (Hart, 2006).

We therefore ran the two-way ANOVA on each subscale measurement of the NASA-TLX questionnaire results. We screened the results for the same assumptions as before, and eliminated the Temporal Demand subscale as it did not meet the Homogeneity of Variance assumption. The remaining subscales of Effort, Frustration and Mental Demand were not significant between 2D and VR mediums. However, there was a significant difference for Performance, $F(1,\ 12) = 13.816$, $p = 0.003$, and also for Physical Demand, $F(1,\ 12) = 10.026$, $p = 0.008$.

Simple main effects analysis therefore shows that there was a significant difference between Physical Demand in 2D ($M = 8.13$, $SD = 5.94$) and VR ($M = 33.75$ $SD = 21.21$) in our study. The physical activities of movement and body rotation in VR logically support this result. Similarly, there was a significant difference between the participant's reported Performance in 2D ($M = 54.38$, $SD = 16.13$) and in VR ($M = 28.75$, $SD = 10.26$). This result suggests that participants both felt more successful, and more satisfied, after completing data exploration tasks in VR than in 2D.



The complete set of results for this hypothesis is available in Appendix K.2 (including raw results, tests for normality, and Levene's test for equality).

7.3 Insight Results

7.3.1 Testing EH2: Number of Insights

Our second hypothesis predicted that there would be a significant difference between the number of insights generated by participants in the VR visualisations and equivalent 2D visualisations. This hypothesis was prompted by RQ2, which looked to compare the insight generation process in both exploration mediums. Similarly to our process for EH1, we combined participants across both visualisations and evaluated EH2 through a two-way ANOVA.

The assumption of independence holds for this hypothesis and all remaining hypotheses. We tested for data normality for EH2 using the Shapiro-Wilk test (Appendix K.3.1). p values for each group were greater than 0.05 which suggested that the data was normally distributed. Furthermore, Levene's Test for Equality was not significant for the number of insights generated by participants (Appendix K.3.2). We therefore concluded that all assumptions were met.

A two-way ANOVA was conducted to investigate the effect of exploration medium and visualisation on the number of insights generated by participants. There was a statistically significant interaction between exploration medium and visualisation on the number of insights, $F(1,\ 12) = 6.387$, $p = 0.027$. There was a significant difference by visualisation, $F(1,\ 12) = 71.461$, p <0.001, but there was not a statistically significant difference by exploration medium, $F(1,\ 12) = 0.919$, $p = 0.357$. Hence, we failed to reject NH2.

We concluded that there was no significant difference between the number of insights generated by participants in 2D ($M = 18.63$ $SD = 8.60$) and VR ($M = 17.25$ $SD = 5.63$) during our study. There was a significant difference in the number of insights between Be The Data ($M = 11.84$ $SD = 2.03$) and Parallel Planes ($M = 24.00$ $SD = 4.31$), which suggested that participants were able to make a greater number of observations about the dataset with the Parallel Planes visualisations. However, this was in the context of the main interaction between exploration medium and visualisation which meant that we were unable to draw any definitive conclusions. Partial eta-squared values for exploration medium ($\eta_p^2 = 0.357$), visualisation ($\eta_p^2 = 0.856$), and their interaction ($\eta_p^2 = 0.347$)), gives an indication of the large effect size of visualisation and the number of insights in our restricted population size.

7.3.2 Testing EH3: Number of Correct Insights

Our third hypothesis predicted that there would be a significant difference between the number of correct insights generated by participants in the VR visualisations and equivalent 2D visualisations. As with EH1 and EH2, we combined participants across both visualisations and evaluated EH3 with a two-way ANOVA. The data met all assumptions for normality and homogeneity of variance (Appendix K.3.1, K.3.2).

There was not a statistically significant interaction between exploration medium and visualisation on the number of correct insights, $F(1,\ 12) = 3.309$, $p = 0.107$. Simple main effects analysis showed that there was also not a significant difference on exploration medium for our hypothesis EH3, $F(1,\ 12) = 3.488$, $p = 0.086$. Hence, we failed to reject NH3. Nevertheless, there was a significant difference for the number of correct insights between Be The Data ($M = 6.38$ $SD = 4.50$) and Parallel Planes ($M = 17.88$ $SD = 4.73$), $F(1,\ 12) = 32.806$, p <0.001.

Following the suggestion from EH2 that participants will make a greater number of insights in the Parallel Planes visualisation, the statistical significance of a greater number of correct insights in this visualisation as well is perhaps a rational conclusion. In order to account for disparity in the underlying insight population between groups, Figure 7.4 shows the relative frequency of correct, neutral, and incorrect insights. The number of insights was 149, 138, 95 and 192 for the combined groups of 2D, VR, Be The Data, and Parallel Planes respectively.

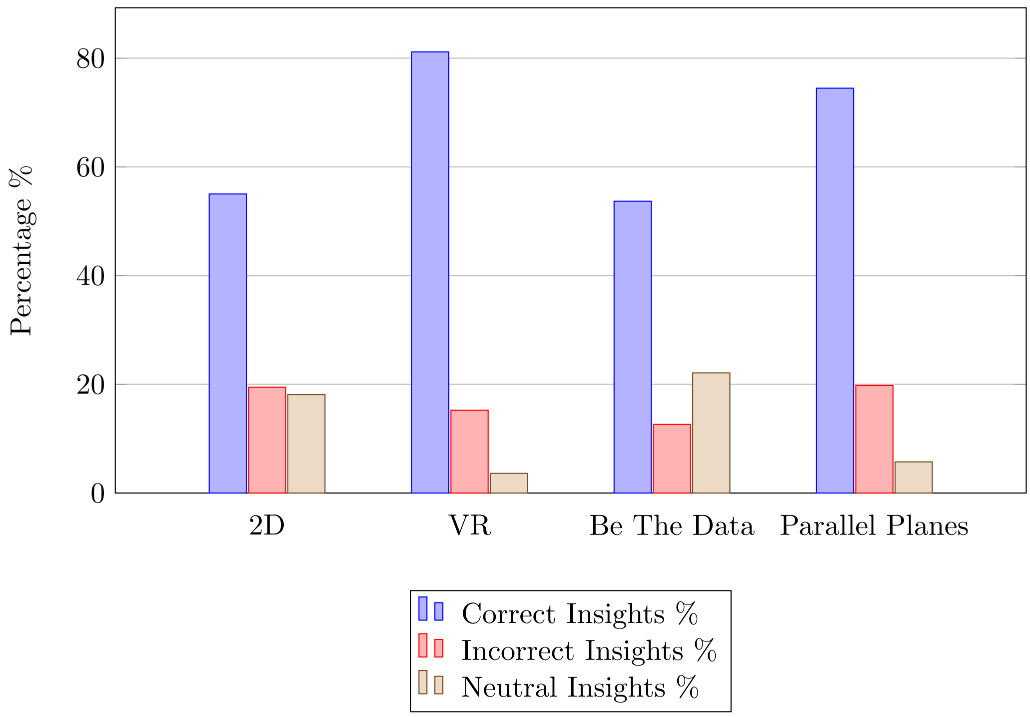


Figure 7.4: Relative Frequency Of Correctness Insights

As previously stated, the number of correct insights between the combined 2D and VR groups was not statistically significant ($p = 0.086$) and as a result we failed to reject NH3. However, focusing purely on correct insights does not give an indication of the number of *incorrect* insights, as we have added a measure for *neutral* insights on our nominal scale. We report that our study suggests that participants have fewer incorrect insights in VR ($M = 0.63$ $SD = 0.52$) compared to 2D ($M = 3.38$ $SD = 1.85$). Furthermore, participants were more likely to make an incorrect observation in the Parallel Planes visualisation ($M = 1.38$ $SD = 1.19$) than in the Be The Data visualisation ($M = 2.63$ $SD = 2.39$). This last statement should be interpreted with the consideration that the visualisations displayed different datasets.

7.3.3 Testing EH4: Number of Unexpected Insights

Our penultimate hypothesis predicted that there would be a significant difference between the number of unexpected insights generated by participants in the Be The Data visualisation and Parallel Planes visualisation. The data met assumptions for normality and homogeneity of variance (Appendix K.3.1, K.3.2) and so we evaluated EH4 using a two-way ANOVA.

There was not a statistically significant interaction between exploration medium and visualisation on the number of unexpected insights, nor was there a significant difference for the exploration medium, $F(1, 12) = 0.753$, $p = 0.403$. There was a statistically significant difference for the visualisation main effect, $F(1, 12) = 26.610$, $p < 0.001$, and consequently we rejected NH4.

Participants were more likely to deviate from their initial exploration questions in the Parallel Planes visualisation ($M = 9.00$ $SD = 4.78$), than in the Be The Data visualisation ($M = 0.25$ $SD = 0.463$). This suggests the additional dimensionality of the Parallel Planes visualisation leads to participants discovering observations without explicitly searching for them. It may also imply that the number of new hypotheses generated by participants is higher during this visualisation. This will be verified during the analysis of EH5.

With results to support EH4, we analysed participant's initial exploration questions to supplement our findings. A frequency analysis of questions is included in Tables 7.1 and 7.2. For the Be The Data visualisation, every participant had at least one question relating to the correlation between sleep and productivity. Other questions for this visualisation were more specific, and often only considered one dimension.

In contrast, the dimensionality of the Parallel Planes visualisation allowed participants to consider a larger number of dimensions in their initial exploratory questions. For the majority of questions participants chose to compare a simple correlation of two dimensions, rather than devising deeper, more specific questions similar to the ones for the Be The Data visualisation.

This may have had some bearing on the number of unexpected insights between visualisations. Anecdotally, we had observed participants struggling to form additional questions for the Be The Data visualisation after writing down *'Correlation between Sleep and Productivity'*. For the Parallel Planes visualisation, participants picked combinations of the five dimensions with relative ease. One possible interpretation of EH4 is that participants who had spent greater effort deliberating over initial questions may have been more focused on answering them. Conversely, participants who were not as attached to their initial questions were more likely to deviate from them, resulting in a greater number of unexpected insights.

Questions for Be The Data	/24
Correlation between Sleep and Productivity	9
Does certain Sleep one day mean certain Productivity the next day?	4
Is there a pattern in Productivity/Sleep?	4
Is there an optimum level of Sleep for Productivity?	2
Does the weekend affect either Sleep or Productivity?	1
Do you Sleep more on certain days?	1
Do you Sleep better after Productive days?	1
Deviance of Sleep and Productivity	1
Comparing personal Sleep habits	1

Table 7.1: Exploration questions for the Be The Data visualisation

Questions for Parallel Planes	**/24**
Correlation between Sleep and Productivity	5
Correlation between Sleep and Mood	4
Correlation between Sleep and Steps	3
Correlation between Productivity and Mood	2
Correlation between Productivity and Tracks	2
Correlation between Steps and Mood	2
Does Mood/Productivity change throughout the week?	2
Correlation between Tracks and Mood	1
Correlation between Productivity and Steps	1
Correlation between Sleep and all others	1
Does the data change throughout the week?	1

Table 7.2: Exploration questions for the Parallel Planes visualisation

7.3.4 Testing EH5: Number of Hypotheses

Our final hypothesis predicted that there would be a significant difference between the number of hypotheses generated by participants in the Be The Data visualisation and Parallel Planes visualisation. We first screened the data for violations of homoscedasticity and normality, but found that the data was not normally distributed for the Be The Data visualisation.

We looked to normalise the number of hypotheses over the entire dataset, but the inclusion of zero-data in the Be The Data groups made this challenging. Box and Cox (1964) cite the Kleczkowski transform suitable for normalisation as follows:

$$z = log(x + c)$$

where c is an arbitrary constant. We applied this transform over our data with $c = 0.5$, half the minimum observation value, rechecked the assumption for homogeneity of variance and then ran the two-way ANOVA.

Given that the transform we applied fundamentally changed our data, the statistical significance of the two-way ANOVA should be interpreted with caution. Nevertheless, there was not a statistically significant interaction between exploration medium and visualisation on the number of hypotheses, nor was there a significant difference for the exploration medium, $F(1, 12) = 0.046$, $p = 0.834$. There was a statistically significant different for the visualisation, $F(1, 12) = 36.891$, $p < 0.001$. Consequently we rejected NH5.

The results support EH5, which suggests that the Parallel Planes visualisation (M = 5 SD = 1.927) encouraged participants to both create and answer questions of the data. Comparatively, the Be The Data visualisation (M = 0.5 SD = 0.926) was not as successful at supporting the creation of hypotheses for participants.

Relation To Depth-Based Insights

We were interested in seeing if an increase in hypotheses also corresponded to an increase in the depths of observations made between visualisations. The act of questioning the data could perhaps imply a deep level of understanding about the data and its characteristics. After checking that the assumptions of normality and homogeneity of variance were met, we ran a follow-up two-way ANOVA on insights labelled as *depth* insights.

There was a statistically significant interaction between exploration medium and visualisation on the number of depth insights, $F(1, 12) = 18.151$, $p = 0.001$. There was also a statistically significant difference on both main effects of medium, $F(1, 12) = 4.856$, $p = 0.048$, and the visualisation, $F(1, 12) = 121.403$, $p < 0.001$.

These results suggest that, as well generating a larger number of hypotheses, the insights which participants had in the Parallel Planes visualisation were deeper (M = 11.75 SD = 3.495) in comparison to the Be The Data visualisation (M = 2.38 SD = 1.506). They also suggest that the use of VR did not result in as many depth-based insights (M = 6.13 SD = 3.482) as the traditional 2D visualisations (M = 8.00 SD = 7.111).

One possible interpretation of the difference between the Be The Data and Parallel Planes visualisations is that participants were more comfortable questioning the dataset once they had a deeper understanding of it. The multidimensionality of the Parallel Planes visualisation and its interactivity enabled participants to pick out small subsets of the distribution and reason about their shape. We had assumed that the selection of individual data points in the Be The Data visualisation would support depth-based reasoning as well, but in fact breadth-based insights significantly outweighed depth-based insights for all participants in traditional 2D and in VR (Appendix K.3.3).

Relation To Insights With Valuable Domain Values

During our data analysis we also labelled insights by domain values which we report through Figures 7.5 and 7.6. This allowed us to interpret the number of hypotheses measured by EH5 in the context of more valuable insights with domain values such as DV3 and DV4 (*DV 1*: General Observation; *DV 2*: Weighted Observation; *DV 3*: Explained Observation; *DV 4*: Explicit Hypothesis).

When testing EH2, we had found that there was not a significant difference between the number of insights between mediums, but there was a significant difference between visualisations in the context of the interaction. Figure 7.5 now visualises this clearly with insights broken down by their domain values. Figure 7.6 represents the domain values of insights as a percentage of the total number of insights for the particular visualisation, or medium.

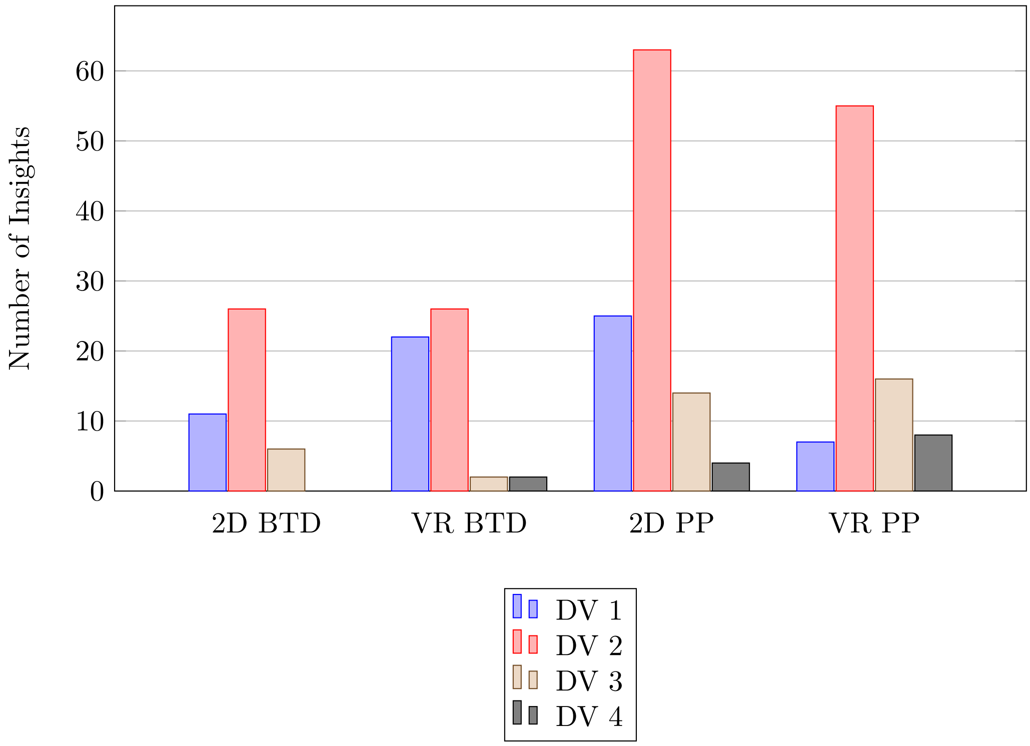


Figure 7.5: Frequency Of Domain Values Between Each Experiment

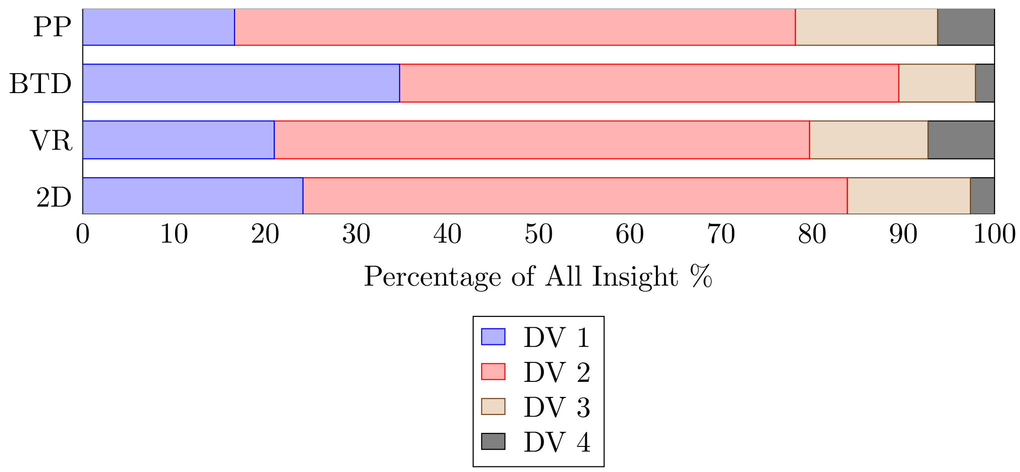


Figure 7.6: Domain Value Breakdown Between Each Visualisation And Medium

The traditional 2D medium did result in a larger, statistically significant number of depth-based insights than VR, but there was no significant difference between the number of hypotheses created between exploration mediums. However, in terms of domain values VR had a larger number of DV4 insights. This suggests that participants established a greater number of explicit hypotheses in VR than in 2D. In contrast, there were a greater number of hypotheses which were deemed to be part of the DV3 category in 2D. For instance, a DV3 hypothetical insight of *"On Thursday [tracks] peaks again, so maybe they need that full productivity for all of those tracks"* is not as valuable as DV4 insight *"Sleep isn't really related to mood, they seem tangential – I wonder if sleep and steps are correlated?"*. Nevertheless, the domain value results in particular should be interpreted cautiously due to the subjective way in which DV3 and DV4 insights can be categorised.

The suggestion from these results is that the Parallel Planes visualisation supported a far greater depth-based exploration of the dataset, enabling participants to interpret more significant and valuable insights. Non-expert participants understood the visualisation and were comfortable brushing subsets of the data, allowing them to answer and create new hypotheses. In comparison, the Be The Data visualisation was not as successful with hypotheses generation. We suggest that this may be due to the visualisation resulting in more breadth-based insights, rather than depth-based insights.

The following section supplements our findings for the experimental hypotheses with comments from study participants which occurred during and after the study.

7.4 Participant Feedback

We collected a diverse set of comments during the study. Participants were most enthused by the Parallel Planes visualisation, whereas feedback was more mixed for the Be The Data visualisation. This section will summarise the specific feedback we received around each visualisation, as well as outlining more general comments on the exploratory data process.

7.4.1 Parallel Planes

Common feedback for the Parallel Planes visualisation was centered around the number of lines between the planes. One participant commented that most lines *"weren't relevant"* and that if they had to *"simplify it after looking at it, you might as well simplify it before delivering it"*. A separate participant expanded upon this by stating that *"it would be easier if you had an average line for each day"*. Our observations of participants gravitating towards selecting outliers in the Parallel Planes visualisation emphasise these comments. Participants typically focused on edges cases, rather than lines in an average grouping for each day.

This feedback indicates that participants may have felt overwhelmed by the amount of lines along the planes that was presented to them, or that there was additional cognitive effort required to extract the meaningful insights they were interested in. This is consistent with findings by Choe et al. (2014) and Jones and Kelly (2016).

Choe et al. suggest that variables that do not seem to correlate with anything should be removed early in the stage-based model of personal informatics. The feedback and behaviour of participants in our study suggests an alternative approach. The absence of correlation in the Parallel Planes visualisation was often an area of interest – participants were interested in uncorrelated variables, as well as the distribution of an average day across the planes. However, participants were not interested in the observations between the average and the outliers. We anticipate that reducing the number of lines to a single average case, and retaining specific outliers, would enhance the insight-generation process and reduce the cognitive effort required to interpret the visualisation.

The highlighting interaction received positive feedback. Participants commented that it was intuitive to understand and essential for detecting the shape of a day across different dimensions in the visualisation – both in the traditional 2D visualisation and in VR. In terms of dimensionality, the multiple parallel coordinate visualisation was complex for participants to comprehend. Several participants commented on the difficulty of interpreting the webpage with multiple parallel coordinates: *"Having graphs for different days felt quite disjointed ... navigating between them and recalling what I was last on was a little challenging"*.

The additional z axis in VR meant that the Parallel Planes visualisation did not have this challenge. The visualisation was not disjointed – participants were able to interpret the dataset across the recorded dimensions in a day, or in the z axis along the weekdays. The visualisation supported participants to interpret a previously disconnected dataset in a combined manner. One participant commented that "*having the multidimensional stuff is great. It forces you think think through things a bit more in terms of causality*". The nature of the visualisation in VR therefore meant that participants were able to obtain a holistic overview of the dataset through a single visualisation.

7.4.2 Be The Data

We transformed the concept of becoming a data point into a VR environment from the initial work of Chen et al. (2016). Yet, in our study, participants generally did not use this functionality as part of their main data exploration process. One participant commented that they preferred to be positioned at *"the end or in the middle, rather than being at a specific point"* in the graph. In fact, we had observed all participants making 2D projections of the data from ends of the graph. This form of orthographic projection reduced the dimensionality of the three-dimensional graph into two dimensions, which resulted in the number of breadth-based insights far outweighing depth-based insights.

One participant offered a potential reason for this: *"We were always taught at school not to use 3D charts... you're not used to analysing data in a 3D way"*. Another participant explicitly noticed their behaviour and commented *"I feel like it almost defeats the point ... I have a 3D dataset and one of the first things I'm drawn to do is look at it from one-end, or top-down, to get different 2D sets"*. It was therefore not clear whether the use of VR and a three-dimensional chart offered many benefits over a standard two-dimensional graph in terms of the participant's data understanding. Nevertheless, participants were engaged during the data exploration. This could be due to participants having to focus harder to interpret the dataset in three-dimensions, the additional interaction techniques provided by VR, or merely because the visualisation and VR was a new experience. Ultimately, the novel VR visualisation did engross participants, and on average scored better in terms of frustration and perceived performance in task workload subscales comparative to its 2D equivalent (Appendix K.6).

Feedback we received for this visualisation emphasised that the prototyping process only informed the direction of design decisions, and was not able to capture the level of feedback we received in the main study. For instance, in both the 2D and VR implementations of Be The Data, participants commented that displaying the date in a format such as '29-08-16' was not as useful as seeing 'Monday'. The abstraction of the dataset was a common theme in both Be The Data and Parallel Planes visualisations – participants wanted to see further detail about data points or lines when they had been selected, but they did not want this information to be available immediately. This form of feedback shows the importance of conducting wider studies to capture additional observations from participants. Given the demographic in this study ($N = 16$, 13/16 with A/A$*$ in A Level Mathematics), until the visualisation is released to a wider and more diverse population, it is not possible to obtain representative opinions on the design and functionality of both visualisations.

7.4.3 Inital Exploration Questions

In section 7.3.3 we touched upon the difficulty which participants had formulating questions to ask of the dataset, particularly when the data had a smaller number of dimensions. Initially this was not unsurprising. A limitation in our study was that participants were not interacting with their own tracked data, and so may not have been able to compose questions around the dimensions they were interested in. In spite of this, we observed participants constructing questions with relative ease for the Parallel Planes visualisation. We discussed how this may have resulted in a greater number of undirected insights in section 7.3.3.

Participants did find it challenging to construct questions for the Be The Data visualisation – we believe that this was due to the reduction in dimensionality. In the Parallel Planes visualisation, there were an additional three dimensions, causing one participant to proclaim: *"You could sit here for hours talking about this graph but I don't really want to!"*. The greater dimensionality led to a large amount of potential insight in this visualisation, which we suspect made it easier for participants to devise questions during the initial exploratory phase . Our findings therefore imply an open challenge for participants tracking a smaller number of dimensions – how can personal informatic systems best frame a self-tracked dataset with a smaller number of dimensions to support and encourage reflection?

7.4.4 Presence And Realism

Given that both visualisations did not perform well for realism, participant feedback concerning this is of particular interest to us. In section 7.1 we discussed how realism was one of three factors of our questionnaire's measurement of presence. Indeed, realism is often considered to be an integral part of presence across the literature (Welch et al., 1996; Khanna et al., 2006). However, to the best of our knowledge, presence has not been measured previously for data visualisation – a domain which realism typically does not lend itself to. This was reflected in the design of our visualisations, and through the low ratings of realism recorded through the IPQ questionnaires.

The implications arising from this are two-fold. Firstly, presence is often used as a success criteria for the evaluation of VR systems – in Chapter 2 we quoted a study where presence was the "*defining factor in the success*" of their system (Hodges et al., 1994, p.10). Is presence therefore an appropriate method for evaluating immersive data visualisations?

Through our experimental hypotheses we explored how the data exploration process was comparatively similar between mediums, with the exception of further analysis in EH5 and individual subscales for task workload. This data exploration process followed one of the key purposes of visualisation – insight – which is more representative of the purpose in which participants used these visualisations. We therefore argue that the measurement of presence is perhaps not as applicable in the data visualisation domain, as it is to other domains such as exposure therapy. If the purpose of visualisation is insight, then presence does not represent an evaluative measure for this, nor does it represent a defining factor of success.

Secondly, is realism necessary for data visualisation? Our presence questionnaire broke the factor of realism down into a behavioural form (Q12: *How much did your experience in the virtual environment seem consistent with your real world experience?*) and a visual form (Q13: *How real did the virtual world seem to you?*). Both of these questions had varied, but generally low ratings for realism which suggests that in our study there was little difference between behavioural and visual realism when evaluated for data visualisation.

Participants performed similarly in VR compared to the traditional 2D visualisations, suggesting that neither behavioural nor visual realism was required for participants to make a success of the data exploration process in Parallel Planes. One participant commented that *"a scene would not have helped with doing the task... (no scene) makes you focus on the data without any distractions"*. This indicates that realism is perhaps not required in order for participants to perform well in the VR data visualisation.

Nevertheless, the same participant also concluded that the lack of a scene in the Parallel Planes visualisation *"doesn't necessarily make it more fun"*. This hints at the question of habituation – something which we were not able to answer in our limited scope study. Although the participant had concluded that a scene would not have helped them with the data exploration process, they stated that *"in terms of should I use VR again, then a scene probably would"*. This is particularly pertinent, given the key abandonment challenges faced by personal informatics discussed in Chapter 2. For commercial smartphone applications, whose business models are often defined by their ability to retain users, this indicates that a scene may be necessary in VR to keep users engaged. Further research is needed to pinpoint whether a scene should be visually realistic, or the visualisation behaviourally realistic, in order to balance preventing abandonment and presenting an interpretable data visualisation.

7.5 Evaluation Summary

This chapter has covered the three methods we used to evaluate our system: the IPQ questionnaire for presence, the NASA-TLX questionnaire for task workload, and Saraiya et al. (2005)'s insight based methodology.

We reported the results for the presence questionnaire, finding that our visualisations performed well in terms of spatial presence and involvement. For the third factor of presence – realism – our visualisations were clearly limited. Nevertheless, we were able to make a comparison to other VR systems in the literature, concluding that both of our visualisations performed favourably in terms of presence.

We failed to reject NH1; that there would not be a difference between participant's task workload in 2D and in VR. However, multiple two-way ANOVAs on each subscale measurement of the TLX-scale yielded two statistically significant differences between physical demand and performance in 2D and VR. These results suggested that there was not a difference in the overall task workload for both exploration mediums, but that participants felt more successful, and more satisfied, after the data exploration process in VR.

There was not a statistically significant difference between the number of insights generated in 2D and VR and hence we failed to reject NH2. However, main effects analysis determined a significant difference between the number of insights generated between the two visualisations – the Parallel Planes visualisation was the most powerful in terms of number of insights generated. With more than double the number of insights on average, this visualisation clearly supported participants to reflect on a high-dimensionality dataset.

Our next hypothesis, EH3, focused on the correctness of these insights. There was not a significant difference between 2D and VR, and we therefore failed to reject NH3. Nevertheless there was a significant difference between the number of correct insights between visualisations. With the disparity in frequency of insights accounted for between visualisations, we reported that participants were more likely to make a correct insight in the Parallel Planes visualisation than in the Be The Data visualisation.

We found evidence to support EH4 (p <0.001, $\alpha = 0.5$); that there would be a significant difference between the number of unexpected insights in each visualisation. Participants were more likely to deviate from their initial questions in the Parallel Planes visualisation ($M = 9.00$ $SD = 4.78$), than in the Be The Data visualisation ($M = 0.25$ $SD = 0.463$). An analysis of initial exploration questions followed, where we suggested that the ease at which participants created questions for the Parallel Planes visualisation with a multidimensional dataset led to them being less attached to their initial questions.

We also found evidence to support our final hypothesis (p <0.001, $\alpha = 0.5$), which predicted that there would be a difference between the number of hypotheses generated between visualisations. Analysis followed in which we discussed the relation to depth-based insights and insights with valuable domain values. Our results suggested that the Parallel Planes visualisation supported a depth-based exploration of the dataset, enabling participants to ask and answer questions, and interpret more valuable insights.

The chapter concluded with a discussion of participant's informal feedback with regards to both visualisations. The strengths and limitations of the visualisations, as well as initial exploration questions and realism, were placed in the context of feedback from participants.

The next chapter will define the contribution of our visualisations and evaluation, and formally detail the limitations to our study. We finally outline extensions to this work in the form of future research directions.

Chapter 8

Discussion and Conclusion

The discussion continues in this chapter, with the implications of our results discussed in the context of research question outcomes. Limitations of our empirical evaluation are reviewed, and we make suggestions for future research which adopts Saraiya et al.'s insight-based methodology as an evaluation approach. Finally, the contribution of this dissertation is summarised, and we outline future research directions for the *Be The Data* and *Parallel Planes* visualisation techniques

8.1 Research Question Outcomes

In Chapter 6 we formalised the three following research questions:

- **RQ1:** Is the perceived workload for data exploration in VR different to exploration using traditional paradigms?
- **RQ2:** Does the use of VR in either *Be The Data* or *Parallel Planes* visualisations affect the insight generation process?
- **RQ3:** Does the *Parallel Planes* visualisation support the interpretation of a high-dimension dataset for non-expert users using VR?

Through our empirical evaluation we found evidence that supported a selection of our experimental hypotheses, and evidence which meant we were failed to reject certain null hypotheses. Multiple hypotheses were derived from our research questions, allowing us to answer these research questions using a variety of metrics. The following sections will communicate the outcome of each research question.

RQ1

For our first research question, we measured participant's perceived task workload through the NASA task workload questionnaire. The same dataset was compared between a traditional 2D paradigm and a VR visualisation developed for the Google Daydream platform. The headline outcome for RQ1 was that there were not a significant difference for task workload between data exploration in a traditional visualisation and in VR.

This in itself was a notable finding, because it suggested that there was no reduction or increase in overall task workload, despite the clear visual and functional distinctions between traditional and VR paradigms. Our evaluation therefore focused on analysing workload at a greater granularity of individual workload subscales.

We found a significant difference between physical demand in 2D and VR, logically showing that the additional physical movement and rotation in VR had a significant effect on this subscale. We also found that participants felt more successful and satisfied after completing the data exploration in VR than in 2D. These results have interesting implications, suggesting that although participants had to work physically harder, they felt more accomplished after completing the data exploration in VR.

Section 8.2 places these granular findings into the context of potential limitations, such as the novelty of the VR visualisation on participant's perceived workload. Nevertheless, we did not find a significant difference for overall task workload for data exploration between VR and traditional 2D paradigms.

RQ2

RQ2 questioned whether the use of VR in either visualisation would affect the insight generation process. Through our insight-based methodology, four hypotheses were tested with multiple two-way ANOVAs. We found several differences in the insight generation process, but suggest that this question would benefit from further research due to the differences between the Be The Data and Parallel Planes visualisations.

In EH2 and EH3, we did not find a significant difference between the number of insights and the number of correct insights generated in 2D and VR respectively. Similarly, there was not a significant difference between the number of unexpected insights and the number of hypotheses generated in 2D and VR (EH4, EH5).

Nevertheless in our study we reported that participants were more likely to have fewer incorrect insights, and fewer neutral insights, in VR than in 2D. Furthermore, our results indicated that participants obtained more valuable insights in VR, though neither of these claims were statistically proven. Finally, and conversely, we found a significant difference between the number of depth-based insights, suggesting that participants made more depth-based observations over the two visualisations in the traditional 2D format.

Due to disparity in the final insight metrics for each visualisation, and the limited sample size in our study, we suggest the need for a larger study which does not combine visualisations during analysis, and enables a single visualisation to be fully compared between exploration mediums.

Based on the measures of insight in our hypotheses, there was a difference in the insight generation process according to the exploration medium used. However, only one of these differences was statistically proven, and so we are unable to give a definitive answer for RQ2. We therefore report these results, which illustrate the similarities and differences on the insight generation process between exploration mediums.

RQ3

Our final research question explored whether the Parallel Planes visualisation supported users to interpret a highly-dimensional dataset. Our findings in EH2, EH3, EH4 and EH5 imply that the Parallel Planes visualisation was a successful technique for the exploration of a multi-faceted dataset for non-expert users.

During our empirical evaluation, we found statistically significant results for the Parallel Planes visualisation relating to the number of: insights, correct insights, unexpected insights, depth-based insights and hypotheses. Furthermore, we reported that participants were able to extract more valuable insights using the Parallel Planes visualisation.

The results collected during this study strongly suggest that the Parallel Planes visualisation enabled users to interpret a multidimensional dataset successfully. The visualisation chained together separate dimensions of the dataset, enabling participants to view a comprehensive representation of connections between multiple variables. Participants arrived at significant discoveries by conducting a depth-based exploration of the visualisation, selecting subsets of the dataset and reasoning about it. Our evaluative approach captured this behaviour, and indicates that the Parallel Planes visualisation technique positively supported users to interpret a multidimensional dataset.

8.2 Limitations

There were multiple limitations to this project which we cover over the two following sections. Firstly, the scope of empirical evaluation was restricted, and we did not consider the data reflection process as part of a wider, more-detailed study. Secondly, Saraiya et al.'s insight-based methodology was extensive, and we make suggestions for improving the scope and accuracy of this approach in the future. After reviewing limitations, we then detail the contributions of this dissertation.

8.2.1 Scope of Empirical Evaluation

One of the main limitations of this study was the limited scope of the empirical evaluation. The use of these visualisations was considered only in isolation by participants – purely as mediums for reflecting on personal data. Li et al. (2010) show this to be just one of five stages of the personal informatics model, which our study exclusively evaluated. A longer-term study would assess this model with an end-to-end process to see if there are implications for goal-setting and behavioural change after data exploration in VR.

Indeed, a longer term study would account for participants both collecting and reflecting on their own tracked data, rather than reflecting on the tracked data of others. While this was beneficial for allowing us to fairly compare participants between experiments, we anticipate that the insight generation process could change due to the additional context which participants may be able to bring to a real dataset of their own.

Additionally, our evaluation did not consider habituation, and how this could affect the long-term use of the visualisations and VR in supporting reflection of personal data. In a longer-term study we would therefore explore the effects on measures such as task workload. For instance, our evaluation found that participants felt more successful and satisfied after data exploration in VR. However, it is not clear whether the novelty of completing this task with an unfamiliar technology had any bearing on this perceived measure of workload, and whether similar participant responses would remain consistent or decrease over time.

Finally the participants in the study must be considered. We used a randomised block design to distribute participants by gender over experiments, achieving a relatively fair balance of two groups (8 Female, 7 Male + 1 Prefer Not To Say). Nevertheless, this design did not consider the demographic of participants – all undergraduates at the University of Bath ($M = 21.6$), and the majority (13/16) with Grade A$*$ or A equivalent at A Level Mathematics. The mathematical background of our participants should be considered when interpreting the contributions of our project as their backgrounds may indicate a stronger ability to interpret the visualisations than a sample from a wider demographic. This is a limitation which could be resolved through further evaluation, with participants selected from a diverse range of technical and non-technical backgrounds.

8.2.2 Insight-Based Evaluation Methodology

We adopted Saraiya et al. (2005)'s insight-based approach for evaluating our visualisations in VR. While this methodology enabled us to capture an extensive amount of insight-related results, it came with several drawbacks which we summarise and make recommendations for below.

Target Select Criteria

The data analysis process to first extract insights which participants got from the visualisations and then code them was intensive and time-consuming even for our limited size study ($N = 16$). This involved transcribing recordings for each participant, extracting all applicable insight from the transcriptions, the categorisation and coding of 6 attributes for each insight, and then checking the correctness of each insight against the visualisations. Saraiya et al. do note that their methodology is labour intensive, and suggest that self-reporting could be a solution. However, we do not believe that self-reporting would have been able to capture both the extent of information and also the diversity of information analysed in our study.

Instead we suggest that studies looking to adopt this evaluation methodology should reduce the scope of post-transcription categorisation and coding of insights into targeted attributes only. For instance, in our analysis we captured information about the category of each insight and a measure of the participant's perceived total insight for the visualisation. However, these measures were not included in the analysis of results as they were either not relevant to the focus of the hypotheses, or there was not a wide enough range of participants to make firm conclusions about the findings. Therefore, we recommend that studies analyse insights using chosen and applicable criteria, rather than the full evaluation criteria proposed by Saraiya et al.. The categorisation and coding process can be retrospectively revisited if additional analysis is required.

Determine Strict Guidelines Through Preprocessing

The subjectivity of the data analysis process was also a limitation to this study. For instance, the act of determining individual insights – when one insight ends, and another insight begins. In addition, identifying the value of an insight in terms of its domain value was challenging, and the subjectivity of our judgements may have had an impact on the accuracy of our insight findings. We recommend that for future evaluations a pool of insights is taken before the main data analysis begins, and preprocessed in order to determine how insights should be separated.

Furthermore, strict guidelines should be formulated to help guide the data analyst to correctly label insights in terms of the domain value. After our piloting process we identified a schema in section 6.4.2 to address this, but this was still open to interpretation for insights that sat between domain values. A more rigid specification would have supported the analyst's decision making and ensured better accuracy in the results.

Visualise And Verbalise

Furthermore, in terms of accuracy, our evaluation was not able to verify whether participants verbal comments were matching with their actions in the virtual environment. We were able to attain an indication of accuracy by comparing the think-aloud protocol with the participant's forward-facing direction and head orientation in VR. Nevertheless, the use of the think-aloud protocol was the main indicator which we had for each participant, with accuracy determined later during data analysis.

We did consider capturing a screen recording of the participant's exploration in VR, but this functionality caused a reduction in frame rate when we tested this alongside VR. A solution which was not feasible within the resource constraints of this project would have been to cast[1] the smartphone screen to a separate monitor. This would require additional hardware, and a smartphone which could handle both VR and the casting process simultaneously. However, this would let analysts observe the accuracy of participant's observations in real-time, rather than post-experiment. Furthermore, it would allow analysts to record insights which participants may not have verbalised.

Extension Metrics

A significant limitation during our application of Saraiya et al.'s insight-based methodology was that we did not label insights by the time at which they occurred – this was a decision made due to the timing constraints of the project. Future studies could record this attribute and make timing analyses such as *'Average time to first hypotheses', Average time recognise first cluster'*. Timing insights indicate that users get immersed in the dataset more quickly, suggesting a faster visualisation learning time (Saraiya et al., 2005).

A further metric which this study did not measure was the confidence of insights which participants were making. By observation, participants seemed to be less encouraged to think-aloud for the traditional 2D Be The Data visualisation compared to the VR Be The Data visualisation where participants made confident assertions about the dataset. There did not seem to be a difference in our anecdotal observations of the Parallel Planes visualisation between mediums.

This phenomenon may of course be representative of our limited study size, and the demographic randomly assigned to the 2D Be The Data group. However, this could also represent a distinction between mediums, in particular for the lower dimensionality dataset. An extension of the insight-based methodology could consider confidence retrospectively by parsing and rating confidence for key phrases such as *'I think, it seems like, I can definitely see that...'*. Alternatively, participants could be directly asked how confident they were about individual observations made during the experimental process itself. A significant difference in the confidence of insights between mediums could be quantified, with confident observations perhaps implying a deeper level of understanding of the dataset, which may lead to a greater number of hypotheses put forward by the participant.

The following section will summarise our contributions to the personal informatics, visualisation and virtual reality fields.

[1] https://www.google.com/chromecast/built-in/

8.3 Summary of Contributions

During this dissertation we have introduced the first applications of immersive data visualisations to support users to reflect on their collected personal data. These visualisations sit firmly in the reflection stage of Li et al. (2010)'s stage-based model, enabling the discovery of patterns and correlations in datasets of varying dimensionality. Our three-part approach to the empirical evaluation of these visualisations shows particular promise for the Parallel Planes visualisation, which allowed participants to successfully navigate a multi-faceted dataset and extract valuable insights.

The Be The Data visualisation applies the physical concept of becoming a data point by Chen et al. (2016) to a virtual world. Users can fly around in this world, interpreting the three-dimensional scatter plot visualisation from a number of perspectives including the perspective of a data point. Our evaluation found that this experience absorbed participant's attention in terms of the involvement factor of presence, but did not significantly change the resultant insight characteristics in comparison to a traditional 2D visualisation. In fact, our observations of participants indicate that they were making orthographic projections to interpret the three-dimensional VR visualisation in a similar fashion to traditional 2D visualisations.

The Parallel Planes visualisation is an extension of a conventional parallel coordinates visualisation, allowing participants to reflect on multidimensional datasets in an immersive setting. We built on research by Brunhart-Lupo et al. (2016) who added a z axis to each coordinate to transform them into planes in the visualisation. The visualisation integrates a disconnected dataset and enables users to perceive a holistic overview of interrelated dimensions. Our evaluation focused on whether this form of visualisation could be understood by non-experts to gather insights from a multi-faceted personal dataset. We found that participants used the interaction techniques to conduct a depth-based exploration of the visualisation, arriving at compelling and meaningful discoveries.

Our final contribution is the evaluation of both visualisations with regards to presence, task workload, and insight. Significantly, we suggest that presence may not be the most appropriate evaluation method for assessing data visualisations in VR due to its close ties to realism. We also make recommendations for future studies looking to use Saraiya et al. (2005)'s insight-based evaluation approach. Our full results indicate that overall task workload between 2D and VR is similar, but that participants feel more successful and satisfied after exploring datasets in VR. Furthermore, that the Parallel Planes visualisation supports the interpretation of a multi-faceted dataset, enhancing the data reflection process through the discovery of valuable insights, a significant increase in the number of hypotheses, and deeper unexpected observations.

8.4 Future Research Directions

During section 7.4 we described the challenges of evaluating data visualisations using presence. We also explored where realism fits into this process, and whether the inclusion of a scene could lead to longer-term engagement in VR, identifying a research gap in the literature. Alongside section 8.2 which discussed the limitations of Saraiya et al.'s insight-based methodology for evaluation, both of these sections form recommendations for future research into empirical evaluations of data visualisations in VR.

In this final section, we outline theoretical and technical extensions to the two visualisations evaluated during this study.

8.4.1 Be The Data

Predefined Viewpoints

During our study, participants navigated in the virtual environment such that they could interpret the Be The Data visualisation in two-dimensions, through orthographic projections. Several participants commented that they were not sure if their position was centrally aligned on the axes while looking back down the dataset, which may have had an impact on the accuracy of their inferences. Given that we found evidence to support that users in VR felt more satisfied and successful during data exploration than in traditional 2D paradigms, future functionality should further enhance the exploration process to support navigation of the virtual world.

Predefined viewpoints could be visible in the world, both inside the graph, and centrally at either end of the axes. Users could click on these viewpoints and be transported to them, knowing that observations made from these positions would not be skewed or affected by minor variations along the axes which occur during normal swipe-based movement. This would be an example of mixed-initiative interaction discussed in Chapter 2, where system and user agents work together, supporting the discovery and reflection of accurate data insights.

Dynamic And Predicted Data Representation

The Be The Data visualisation displayed a static version of the personal dataset. The dataset was represented using a scatter plot, and the data points in the visualisation did not move during data exploration. One extension to this visualisation would be to reduce the number of days visualised to one week, rather than several months, and allow the user to swipe between weeks from the perspective of a data point.

For instance, from the perspective of a data point situated at 'Monday' in the visualisation, the user could look down the graph and swipe on the controller to move to the next week of data. Variations in the dataset would cause the user to move as the data point within the virtual world, and observe the changes happening in the data surrounding them from week to week.

This could be further extended to partially complete weeks of the dataset, where users view predictions of where their week is heading based on past behaviour. Smartphone applications such as Fitbit could propose goals modelled using historical data, in order to engage users to reflect on their current behaviour and promote positive changes in behaviour through weekly goals.

Triggers For Engagement

Unlike the Parallel Planes visualisation, we do not envisage the Be The Data visualisation as a replacement for existing data reflection tools. Rather, we view it as a complement to existing visualisations presented in dashboards, such as natural language statements like *'You are more productive when you sleep more'*. During Chapter 2 we discussed the separation of high-level and low-level data in the dashboard paradigm. Much like this separation, we anticipate natural language statements as a high-level summary of the visualisation, with a route into virtual worlds where participants can reflect upon this summary in greater detail.

To accompany this method of personal data reflection in smartphone applications, optional notifications could be included. For instance, push notifications to encourage users to reflect on their behaviour in VR at specified intervals. The tension between offering too many notifications and the dataset not changing enough between separate data reflection sessions would have to be balanced. Nevertheless, this could stimulate users to periodically reflect on their personal data, which is particularly applicable to those with goals in mind for behavioural change. Significantly, a single smartphone could be used for collection of personal data and providing the motivation for users to also reflect upon this data in VR.

8.4.2 Parallel Planes

Filtering Observations

In their paper describing the original application of the Parallel Planes visualisation, Brunhart-Lupo et al. (2016, p.1) state that the transformation of parallel coordinates from a surface to a volume "*alleviates the over-plotting problem*". However, for our participants who initially viewed a large quantity of lines across the plane, we found that not all were comfortable interpreting this amount of information. This may be reflective of the different audiences which interacted with each visualisation – Brunhart-Lupo et al.'s research focused on the visualisation's utility with simulation experts.

This finding was consistent with sensemaking challenges identified by Jones and Kelly (2016) for two-dimensional graphical visualisations – in particular, that the quantity of personal information can overwhelm users. The use of VR as the exploration medium did not change this, and we therefore recommend that the amount of lines shown in the Parallel Planes visualisation is filtered initially. Further research is required to determine the extent of filtering. Participant feedback and anecdotal observations of typical behaviour during our study indicate that the average observation for each dimension should be included, as well as a selection of outliers. Choosing the outliers then becomes an elaborate challenge when the lowest and highest observations need to be considered across multiple dimensions of the dataset, and relating to a single observation across all planes.

Data Abstraction

Future work could explore the integration of more granular dimensions into the Parallel Planes visualisation. In our study, one participant wanted more information about sleep because *"the amount of sleep isn't the same as quality of sleep"*. There are opportunities in the visualisation to introduce subdimensions of sleep, such as sleep start and end times, sleep depth, and sleep awakenings. One approach could be to overlay this information once an observation has been selected. Alternatively, the visualisation could be opened up when an observation is selected, and the sleep plane replaced by a series of smaller sub-planes representing these more granular details. This could support users to draw inferences based on definitive low-level data observations, pinpointing them towards specific lifestyle changes they could make.

Extreme Observations And Visualisation Structure

A design constraint discussed during Chapter 4 was that outliers could change the complexion of the Parallel Planes visualisation. Our dataset was normalised to the planes, but extreme values caused tighter clustering of average observations, and a flattened structure which was harder to interpret. This is an open challenge for the parallel coordinates visualisation, and is not just exclusive to the Parallel Planes visualisation evaluated during this study.

One solution in VR could involve restricting the plane to surround only the average cluster of observations in each dimension, or an acceptable range around this average. Outliers could be added in after the construction of the planes, and would be free to leave the boundaries of the plane, observable in VR by looking beneath or above the visualisation. This technique would enable the preservation of the visualisation structure for the majority of observations.

Furthermore, Johansson et al. (2005) explore the use of textures to preserve structure in a parallel coordinates context. They use clustering algorithms to generate textures for the visualisation structure, which distinguish outliers and sub-clusters, supporting analysts to detect patterns across the multiple dimensions. An implementation of this in VR would also reduce the over-plotting challenges previously discussed, but care would need to be taken around the selection of colour, ensuring that appropriate accessibility guidelines were met, and that lines could be distinguished by depth over the z dimension.

The ordering of planes is also an important topic when considering the structure of the multidimensional visualisation. One participant aptly mentioned that the structure of the visualisation *"isn't a feature of the data, it's a feature of the way it's presented"*. To some extent this is true – the ordering of the dataset dimensions can dramatically change connections between planes from horizontal lines to steep angles given the normalised data. An approach to ordering could look towards automatically analysing the dataset to find attractive and interpretable shapes between planes. To supplement this, the development of new interaction techniques in VR could provide additional functionality for manual reordering by users.

Through the Parallel Planes visualisation we discovered that chaining the dimensions in a multi-faceted dataset led users to discover valuable, unexpected insights. The holistic overview of the dataset enabled users to correlate a number of the dimensions, and hypothesise and reflect upon the characteristics of the data. This section has set out future research directions, providing new opportunities for both of our visualisation techniques, and indicates exciting research potential for visualising multidimensional personal datasets in VR moving forward.

Bibliography

Allen, J., Guinn, C. I. and Horvtz, E. (1999), 'Mixed-initiative interaction', *IEEE Intelligent Systems and their Applications* **14**(5), 14–23.

Arms, L., Cook, D. and Cruz-Neira, C. (1999), The benefits of statistical visualization in an immersive environment, *in* 'Virtual Reality, 1999. Proceedings., IEEE', IEEE, pp. 88–95.

Arthur, C. (2015), 'The return of virtual reality: 'this is as big an opportunity as the internet'', *The Guardian* .

Bates, M. (2016), 'Mobile apps study underscores necessity of strong best practices for health and wellness data', `https://fpf.org/2016/08/17/best-practices-consumer-wearables-wellness-apps-devices/`. Accessed: 23rd November 2016.

Bavor, C. (2016), ''(un)folding a virtual journey with google cardboard', `https://blog.google/products/google-vr/unfolding-virtual-journey-cardboard/`. Accessed: 26th January 2016.

Baos, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpi, C. and Alcaniz, M. (2000), 'Presence and reality judgment in virtual environments: a unitary construct?', *CyberPsychology & Behavior* **3**(3), 327–335.

BBC (2016), 'Less than 1 percent of pcs can run virtual reality', `http://www.bbc.co.uk/news/technology-35220974`. Accessed: 27th December 2016.

Bellini, H., Chen, W., Sugiyama, M., Shin, M., Alam, S. and Takayama, D. (2016), 'Virtual & augmented reality understanding the race for the next computing platform'.

Bentley, F. et al. (2013), 'Health mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change', *ACM Transactions on Computer-Human Interaction (TOCHI)* **20**(5), 30.

Box, G. E. and Cox, D. R. (1964), 'An analysis of transformations', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 211–252.

Boyd, D. and Crawford, K. (2011), Six provocations for big data, *in* 'A decade in internet time: Symposium on the dynamics of the internet and society', Oxford Internet Institute, pp. 1–17.

Brandimarte, L., Acquisti, A. and Loewenstein, G. (2013), 'Misplaced confidences privacy and the control paradox', *Social Psychological and Personality Science* **4**(3), 340–347.

Brath, R. and Peters, M. (2004), 'Dashboard design: Why design is important', `http://www.information-management.com/infodirect/20041015/1011285-1.html`. Accessed: 28th November 2016.

Brunhart-Lupo, N., Bush, B. W., Gruchalla, K. and Smith, S. (2016), Simulation exploration through immersive parallel planes: Preprint, Technical report, NREL (National Renewable Energy Laboratory (NREL), Golden, CO (United States)).

Bryson, S. (1996), 'Virtual reality in scientific visualization', *Communications of the ACM* **39**(5), 62–71.

Card, S. K., Mackinlay, J. D. and Shneiderman, B. (1999), *Readings in information visualization: using vision to think*, Morgan Kaufmann.

Cassidy, J. (2016), 'Should we be giving away our personal data for free?', `http://www.bbc.co.uk/news/business-35423513`. Accessed: 22nd November 2016.

Chen, X., Self, J. Z., House, L. and North, C. (2016), 'Be the data: A new approach for immersive analytics'.

Chengzhi, Q., Chenghu, Z. and Tao, P. (2003), Taxonomy of visualization techniques and systems–concerns between users and developers are different.

Chi, E. H.-h. (2000), A taxonomy of visualization techniques using the data state reference model, *in* 'Information Visualization, 2000. InfoVis 2000. IEEE Symposium on', IEEE, pp. 69–75.

Choe, E. K., Lee, B. et al. (2015), 'Characterizing visualization insights from quantified selfers' personal data presentations', *IEEE computer graphics and applications* **35**(4), 28–37.

Choe, E. K. et al. (2014), Understanding quantified-selfers' practices in collecting and exploring personal data, *in* 'Proceedings of the 32nd annual ACM conference on Human factors in computing systems', ACM, pp. 1143–1152.

Cruz-Neira, C., Sandin, D. J. and DeFanti, T. A. (1993), Surround-screen projection-based virtual reality: the design and implementation of the cave, *in* 'Proceedings of the 20th annual conference on Computer graphics and interactive techniques', ACM, pp. 135–142.

Cunningham, J. P. and Ghahramani, Z. (2015), 'Linear dimensionality reduction: Survey, insights, and generalizations', *Journal of Machine Learning Research* **16**, 2859–2900.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K. and Dixon, D. (2011), Gamification. using game-design elements in non-gaming contexts, *in* 'CHI'11 Extended Abstracts on Human Factors in Computing Systems', ACM, pp. 2425–2428.

Dewey, C. (2016), '98 personal data points that facebook uses to target ads to you', https://www.washingtonpost.com/news/the-intersect/wp/2016/08/19/98-personal-data-points-that-facebook-uses-to-target-ads-to-you/. Accessed: 22nd November 2016.

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y. and Mühlberger, A. (2015), 'The impact of perception and presence on emotional reactions: a review of research in virtual reality', *Frontiers in psychology* **6**, 26.

Donalek, C., Djorgovski, S. G., Cioc, A., Wang, A., Zhang, J., Lawler, E., Yeh, S., Mahabal, A., Graham, M., Drake, A. et al. (2014), Immersive and collaborative data visualization using virtual reality platforms, *in* 'Big Data (Big Data), 2014 IEEE International Conference on', IEEE, pp. 609–614.

Economist, T. (2014), 'Hiding from big data: With the increasing commercial use of personal data and multiple security breaches, will people pay for privacy?', http://www.economist.com/news/technology-quarterly/21603233-it-security-increasing-commercial-use-personal-data-and-multiple. Accessed: 22nd November 2016.

Epstein, D. A., Caraway, M., Johnston, C., Ping, A., Fogarty, J. and Munson, S. A. (2016), Beyond abandonment to next steps: Understanding and designing for life after personal informatics tool use, *in* 'Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems', ACM, pp. 1109–1113.

Epstein, D., Cordeiro, F., Bales, E., Fogarty, J. and Munson, S. (2014), Taming data complexity in lifelogs: exploring visual cuts of personal informatics data, *in* 'Proceedings of the 2014 conference on Designing interactive systems', ACM, pp. 667–676.

Ericsson (2015), 'Ericsson mobility report: On the pulse of the networked society', https://www.ericsson.com/res/docs/2016/ericsson-mobility-report-2016.pdf. Accessed: 8th November 2016.

European Commission (2016), 'Regulation (EU) no 2016/679'. (35)
http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679.

Few, S. (2013), *Information Dashboard Design: Displaying data for at-a-glance monitoring*, Analytics Press Burlingame, CA.

Gartner (2013), 'Gartner reveals top predictions for it organizations and users for 2014 and beyond', http://www.gartner.com/newsroom/id/2603215. Accessed: 22nd November 2016.

Gulotta, R., Forlizzi, J., Yang, R. and Newman, M. W. (2016), Fostering engagement with personal informatics systems, *in* 'Proceedings of the 2016 ACM Conference on Designing Interactive Systems', DIS '16, ACM, New York, NY, USA, pp. 286–300.
URL: *http://doi.acm.org/10.1145/2901790.2901803*

Haber, R. B. and McNabb, D. A. (1990), 'Visualization idioms: A conceptual model for scientific visualization systems', *Visualization in scientific computing* **74**, 93.

Haddadi, H. and Brown, I. (2014), 'Quantified self and the privacy challenge'.

Hamari, J., Koivisto, J. and Sarsa, H. (2014), Does gamification work?–a literature review of empirical studies on gamification, *in* '2014 47th Hawaii International Conference on System Sciences', IEEE, pp. 3025–3034.

Hart, S. G. (2006), Nasa-task load index (nasa-tlx); 20 years later, *in* 'Proceedings of the human factors and ergonomics society annual meeting', Vol. 50, Sage Publications Sage CA: Los Angeles, CA, pp. 904–908.

Hart, S. G. and Staveland, L. E. (1988*a*), 'Development of nasa-tlx (task load index): Results of empirical and theoretical research', *Advances in psychology* **52**, 139–183.

Hart, S. G. and Staveland, L. E. (1988*b*), 'Nasa tlx scale', `https://humansystems.arc.nasa.gov/groups/TLX/downloads/TLXScale.pdf"`. Accessed: 2nd March 2017.

Hartson, R. (2003), 'Cognitive, physical, sensory, and functional affordances in interaction design', *Behaviour & Information Technology* **22**(5), 315–338.

Hauser, H., Ledermann, F. and Doleisch, H. (2002), Angular brushing of extended parallel coordinates, *in* 'Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on', IEEE, pp. 127–130.

Heinrich, J. and Weiskopf, D. (2013), 'State of the art of parallel coordinates', *STAR Proceedings of Eurographics* pp. 95–116.

Hendy, K. C., Hamilton, K. M. and Landry, L. N. (1993), 'Measuring subjective workload: when is one scale better than many?', *Human Factors* **35**(4), 579–601.

Henry, J. A. and Polys, N. F. (2010), 'The effects of immersion and navigation on the acquisition of spatial knowledge of abstract data networks', *Procedia Computer Science* **1**(1), 1737–1746.

Hodges, L. F., Kooper, R., Meyer, T. C., De Graaff, J. J. H., Rothbaum, B. O., Opdyke, D., Williford, J. S. and North, M. M. (1994), Presence as the defining factor in a vr application, Technical report, Georgia Institute of Technology.

Hodges, L. F., Kooper, R., Meyer, T. C., Rothbaum, B. O., Opdyke, D., de Graaff, J. J., Williford, J. S. and North, M. M. (1995), 'Virtual environments for treating the fear of heights', *Computer* **28**(7), 27–34.

Horowitz, K. (2004), 'Sega vr: Great idea or wishful thinking?', `http://bit.ly/2h5oHrz`. Accessed: 22nd December 2016.

Horvitz, E. (1999), Principles of mixed-initiative user interfaces, *in* 'Proceedings of the SIGCHI conference on Human Factors in Computing Systems', ACM, pp. 159–166.

Horvitz, E. and Mulligan, D. (2015), 'Data, privacy, and the greater good', *Science* **349**(6245), 253–255.

Hovis, G. (2002), 'Stop searching for information monitor it with dashboard technology', `http://www.information-management.com/infodirect/20020208/4681-1.html`. Accessed: 28th November 2016.

Intille, S. S. (2004), 'A new research challenge: persuasive technology to motivate healthy aging', *IEEE Transactions on information technology in Biomedicine* **8**(3), 235–237.

Jimenez, L. O. and Landgrebe, D. A. (1998), 'Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data', *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **28**(1), 39–54.

Johansson, J., Ljung, P., Jern, M. and Cooper, M. (2005), Revealing structure within clustered parallel coordinates displays, *in* 'Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on', IEEE, pp. 125–132.

Johnson, C. (2004), 'Top scientific visualization research problems', *IEEE Computer Graphics and Applications* **24**(4), 13–17.

Johnson, D. A. and Trivedi, M. M. (2011), Driving style recognition using a smartphone as a sensor platform, *in* '2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)', IEEE, pp. 1609–1615.

Jones, S. and Kelly, R. (2016), 'Sensemaking challenges in personal informatics and self-monitoring systems'.

Jones, S. L. (2015), Exploring correlational information in aggregated quantified self data dashboards, *in* 'Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers', ACM, pp. 1075–1080.

Keim, D. A. (2001), 'Visual exploration of large data sets', *Communications of the ACM* **44**(8), 38–44.

Kenny, R. and Asnes Becker, A. (2015), 'Is the nasdaq in another bubble?: A virtual reality guided tour of 21 years of the nasdaq', `http://graphics.wsj.com/3d-nasdaq/`. Accessed: 22nd January 2017.

Khanna, P., Yu, I., Mortensen, J. and Slater, M. (2006), Presence in response to dynamic visual realism: a preliminary report of an experiment study, *in* 'Proceedings of the ACM symposium on Virtual reality software and technology', ACM, pp. 364–367.

Killock, J. (2016), 'Facebook is right to sink admirals app', `https://www.openrightsgroup.org/blog/2016/facebook-is-right-to-sink-admirals-app`. Accessed: 23rd November 2016.

Lab, G. N. (2016), 'Will the uk brexit?', `http://news-lab-brexit.appspot.com/vr/`. Accessed: 23rd January 2017.

Lanier, J. and Biocca, F. (1992), 'An insider's view of the future of virtual reality', *Journal of communication* **42**(4), 150–172.

Lazar, A., Koehler, C., Tanenbaum, J. and Nguyen, D. H. (2015), Why we use and abandon smart devices, *in* 'Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing', ACM, pp. 635–646.

Lee, M. K., Kim, J., Forlizzi, J. and Kiesler, S. (2015), Personalization revisited: a reflective approach helps people better personalize health services and motivates them to increase physical activity, *in* 'Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing', ACM, pp. 743–754.

Lessiter, J., Freeman, J., Keogh, E. and Davidoff, J. (2001), 'A cross-media presence questionnaire: The itc-sense of presence inventory', *Presence: Teleoperators and virtual environments* **10**(3), 282–297.

Levy, H. P. (2015), 'Wearable technology beyond smartwatches', `http://www.gartner.com/smarterwithgartner/wearable-technology-beyond-smartwatches-3/`. Accessed: 8th November 2016.

Li, I., Dey, A. and Forlizzi, J. (2010), A stage-based model of personal informatics systems, *in* 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', ACM, pp. 557–566.

Li, I., Dey, A. K. and Forlizzi, J. (2011), Understanding my data, myself: supporting self-reflection with ubicomp technologies, *in* 'Proceedings of the 13th international conference on Ubiquitous computing', ACM, pp. 405–414.

Lyons, K. (2016), 'Smartwatch innovation: Exploring a watch-first model', *IEEE Pervasive Computing* **15**(1), 10–13.

Mantovani, F., Castelnuovo, G., Gaggioli, A. and Riva, G. (2003), 'Virtual reality training for health-care professionals', *CyberPsychology & Behavior* **6**(4), 389–395.

Marai, G. E., Forbes, A. G. and Johnson, A. (2016), 'Interdisciplinary immersive analytics at the electronic visualization laboratory: Lessons learned and upcoming challenges'.

Martin, A. R. and Ward, M. O. (1995), High dimensional brushing for interactive exploration of multivariate data, *in* 'Proceedings of the 6th Conference on Visualization'95', IEEE Computer Society, p. 271.

Milgram, P., Takemura, H., Utsumi, A. and Kishino, F. (1995), Augmented reality: A class of displays on the reality-virtuality continuum, *in* 'Photonics for industrial applications', International Society for Optics and Photonics, pp. 282–292.

Mittelstadt, B. D. and Floridi, L. (2016), 'The ethics of big data: current and foreseeable issues in biomedical contexts', *Science and engineering ethics* **22**(2), 303–341.

Moody, G. (2016), 'Going with the flow: The global battle for your personal data', `http://arstechnica.co.uk/tech-policy/2016/11/eu-us-personal-data-flows-explainer/`. Accessed: 22nd November 2016.

Morina, N., Brinkman, W.-P., Hartanto, D. and Emmelkamp, P. M. (2014), 'Sense of presence and anxiety during virtual social interactions between a human and virtual humans', *PeerJ* **2**, e337.

Munzner, T. (2008), Process and pitfalls in writing information visualization research papers, *in* 'Information visualization', Springer, pp. 134–153.

Narayanan, A. and Shmatikov, V. (2008), Robust de-anonymization of large sparse datasets, *in* '2008 IEEE Symposium on Security and Privacy (sp 2008)', IEEE, pp. 111–125.

Neuringer, A. (1981), 'Self-experimentation: A call for change', *Behaviorism* **9**(1), 79–94.

North, C. (2006), 'Toward measuring visualization insight', *IEEE computer graphics and applications* **26**(3), 6–9.

Nunan, D. and Di Domenico, M. (2013), 'Market research and the ethics of big data', *International Journal of Market Research* **55**(4), 2–13.

Ong, S. K. and Nee, A. Y. C. (2013), *Virtual and augmented reality applications in manufacturing*, Springer Science & Business Media.

Pausch, R., Proffitt, D. and Williams, G. (1997), Quantifying immersion in virtual reality, *in* 'Proceedings of the 24th annual conference on Computer graphics and interactive techniques', ACM Press/Addison-Wesley Publishing Co., pp. 13–18.

Payne, S. J. (2003), 'Users' mental models: the very ideas', *HCI models, theories, and frameworks: Toward a multidisciplinary science* pp. 135–156.

Portman, M., Natapov, A. and Fisher-Gewirtzman, D. (2015), 'To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning', *Computers, Environment and Urban Systems* **54**, 376–384.

Pousman, Z., Stasko, J. and Mateas, M. (2007), 'Casual information visualization: Depictions of data in everyday life', *IEEE transactions on visualization and computer graphics* **13**(6), 1145–1152.

Powers, M. B. and Emmelkamp, P. M. (2008), 'Virtual reality exposure therapy for anxiety disorders: A meta-analysis', *Journal of anxiety disorders* **22**(3), 561–569.

Pu, P. and Lalanne, D. (2002), Design visual thinking tools for mixed initiative systems, *in* 'Proceedings of the 7th international conference on Intelligent user interfaces', ACM, pp. 119–126.

Rachuri, K. et al. (2010), Emotionsense: a mobile phones based adaptive platform for experimental social psychology research, *in* 'Proceedings of the 12th ACM international conference on Ubiquitous computing', ACM, pp. 281–290.

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., Gaggioli, A., Botella, C. and Alcañiz, M. (2007), 'Affective interactions using virtual reality: the link between presence and emotions', *CyberPsychology & Behavior* **10**(1), 45–56.

Rooksby, J., Rost, M., Morrison, A. and Chalmers, M. C. (2014), Personal tracking as lived informatics, *in* 'Proceedings of the 32nd annual ACM conference on Human factors in computing systems', ACM, pp. 1163–1172.

Salvendy, G. (2012), *Handbook of human factors and ergonomics*, John Wiley & Sons.

Saraiya, P., North, C. and Duca, K. (2005), 'An insight-based methodology for evaluating bioinformatics visualizations', *IEEE transactions on visualization and computer graphics* **11**(4), 443–456.

Schubert, T., Friedmann, F. and Regenbrecht, H. (2001), 'The experience of presence: Factor analytic insights', *Presence: Teleoperators and virtual environments* **10**(3), 266–281.

Schuemie, M. J. and Van der Mast, C. (1999), Presence: Interacting in virtual reality, *in* 'Proceedings of the 15th Twente Workshop on Language Technology', pp. 213–217.

Schuemie, M. J., Van Der Straaten, P., Krijn, M. and Van Der Mast, C. A. (2001), 'Research on presence in virtual reality: A survey', *CyberPsychology & Behavior* **4**(2), 183–201.

Seymour, N. E., Gallagher, A. G., Roman, S. A., OBrien, M. K., Bansal, V. K., Andersen, D. K. and Satava, R. M. (2002), 'Virtual reality training improves operating room performance: results of a randomized, double-blinded study', *Annals of surgery* **236**(4), 458–464.

Sheridan, T. B. (1992), 'Musings on telepresence and virtual presence', *Presence: Teleoperators & Virtual Environments* **1**(1), 120–126.

Shneiderman, B. and Maes, P. (1997), 'Direct manipulation vs. interface agents', *interactions* **4**(6), 42–61.

Slater, M. and Wilbur, S. (1997), 'A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments', *Presence: Teleoperators and virtual environments* **6**(6), 603–616.

Steed, C. A., Daniel, J. R., Drouhard, M., Hahn, S. E. and Proffen, T. E. (2016), Immersive visual analytics for transformative neutron scattering science, Technical report, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN (United States). Spallation Neutron Source; Oak Ridge Leadership Computing Facility (OLCF).

Steuer, J. (1992), 'Defining virtual reality: Dimensions determining telepresence', *Journal of communication* **42**(4), 73–93.

Stusak, S., Tabard, A., Sauka, F., Khot, R. A. and Butz, A. (2014), 'Activity sculptures: Exploring the impact of physical visualizations on running activity', *IEEE transactions on visualization and computer graphics* **20**(12), 2201–2210.

Tene, O. and Polonetsky, J. (2012), 'Big data for all: Privacy and user control in the age of analytics', *Nw. J. Tech. & Intell. Prop.* **11**, xxvii.

Thierer, A. D. (2015), 'The internet of things and wearable technology: Addressing privacy and security concerns without derailing innovation', *Adam Thierer, The Internet of Things and Wearable Technology: Addressing Privacy and Security Concerns without Derailing Innovation* **21**.

Toffler, A. (1981), *The Third Wave*, Bantam books New York.

Tsai, M.-D., Hsieh, M.-S. and Jou, S.-B. (2001), 'Virtual reality orthopedic surgery simulator', *Computers in biology and medicine* **31**(5), 333–351.

Van Dam, A., Forsberg, A. S., Laidlaw, D. H., LaViola, J. J. and Simpson, R. M. (2000), 'Immersive vr for scientific visualization: A progress report', *IEEE Computer Graphics and Applications* **20**(6), 26–52.

van Gelder, J.-L., Nee, C., Otte, M., Demetriou, A., van Sintemaartensdijk, I. and van Prooijen, J.-W. (2016), 'Virtual burglary exploring the potential of virtual reality to study burglary in action', *Journal of Research in Crime and Delinquency* p. 0022427816663997.

Ware, C. (2004), *Information visualization: perception for design*, Morgan Kaufmann.

Welch, R. B., Blackmon, T. T., Liu, A., Mellers, B. A. and Stark, L. W. (1996), 'The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence', *Presence: Teleoperators & Virtual Environments* **5**(3), 263–273.

Whooley, M., Ploderer, B. and Gray, K. (2014), On the integration of self-tracking data amongst quantified self members, *in* 'Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014-Sand, Sea and Sky-Holiday HCI', BCS, pp. 151–160.

Witmer, B. G. and Singer, M. J. (1998), 'Measuring presence in virtual environments: A presence questionnaire', *Presence: Teleoperators and virtual environments* **7**(3), 225–240.

Zlomuzica, A., Preusser, F., Totzeck, C., Dere, E. and Margraf, J. (2016), 'The impact of different emotional states on the memory for what, where and when features of specific events', *Behavioural brain research* **298**, 181–187.

Appendix A

Initial Requirements List

A.1 Platform Requirements

- FR-1: Users **must** be able to focus on objects and read textual information.
- FR-2: The app **must** maintain head tracking.
- FR-3: The app **must** keep a stable horizon line when a horizon is used.
- FR-4: Camera movement **must** be user-initiated.
- FR-5: The app **must** not interfere with Daydream-level recentering behaviour.
- FR-6: The app **should** recenter with respect to the experience.
- FR-7: The forward direction of app movement **must** be consistent with the Daydream global forward direction.
- FR-8: The app **should** use the (Daydream SDK-provided) neck model.
- FR-9: The app **must** stay in 3D and not display 2D views (e.g Dialogs).
- FR-10: A model of the Daydream controller **must** be rendered in the VR environment.
- FR-11: The Daydream controller **should** be used to click on targets.
- FR-12: Head gaze **could** be used to click on targets.
- FR-13: Swiping on the Daydream controller **must** be used to support movement in the environment.
- FR-14: The app **must** hide the Android system status and navigation bars while in VR.

- FR-15: The app **should** run without crashing to a 2D dialog in VR.
- FR-16: The app **must** maintain high performance and **should** not suffer from thermal throttling.
- FR-17: The app **must** not exhibit characteristics of unexpected distortion when viewed through the Daydream headset.
- FR-18 A loading indicator for the splash screen **could** be added while waiting for the visualisation scenes to load.

- NFR-1: The Google VR SDK for Unity 1.20 **should** be used for development purposes.
- NFR-2: The Unity technical preview for Google Daydream **must** be used. This is currently *5.4.2f-GVR13*.
- NFR-3: The app **must** target Android 7.0 smartphones to enable Daydream platform compatibility. *Note: Daydream is only officially supported on specific Android 7.0 devices.*
- NFR-4: The app **must** maintain high performance across both *Be The Data* and *Parallel Planes* visualisations for at least 15 minutes. High performance is defined as a consistently high frame-rate (60fps).

A.2 Existing System Requirements

- FN-19: The app controls **must** make it easy for a user to move 'inside' the dataset.
- FN-20: There **must** be a smooth transition when moving from an overview of the data to a detailed point inside it.
- FN-21: The app **could** use storytelling to help explain aspects of the dataset.
- FN-22: The user **must** be able to look around in all directions from their vantage point.
- FN-23: The axes labels **must** be spatially separated and must not overlap.
- FN-24: The number of axes labels **should** be reduced where possible. Significant axes labels **must** still be shown.
- FN-25: Textual information **should** be clear and a distinct colour from the data.
- FN-26: Textual information **should** be avoided where possible. Let the data do the speaking.

- FN-27: The user **must** be able to freely move to any position within the data exploration environment.
- FN-28: Sound effects **must** be played when the user clicks on objects within the data exploration environment.
- FN-29: An object which can be selected **should** change colour when clicked by the user.
- FN-30: Extra information about the selected data point **should** overlay the screen when the object is either selected or hovered over.
- FN-31: Visual cues **could** be provided should the user stray too far from the data visualisation.

- NFR-5: It **must** take less than 3 clicks for the user to get 'inside' the dataset from their initial starting point.
- NFR-6: Movement inside the environment **must** follow either a linear or bezier curve.
- NFR-7: A data point's information overlay **should** fit entirely inside the VR stereo rendering. They should not have to move their head from the selection to gather all of the information.

A.3 Be The Data Requirements

- FN-32: The user **must** be able to move outside of the graph and view it from a distance. They must not be constrained by the graph boundaries.
- FN-33: The user **must** be able to move through the axis.
- FN-34: The graph **could** be placed in an environment or scene. e.g In a rolling hill, on top of a skyscraper.
- FN-35: The user **must** be able to hover over a data point and an information overlay will appear, displaying additional information about the current selection.
- FN-36: On clicking the data point, the user **must** be transported to the current position of the selected data point in the environment.
- FN-37: The graph axes **must** be clearly labelled with their main dimension (e.g Sleep) and units along this dimension (e.g 1, 2, 3 hours)
- FN-38: The user **should** be able to move freely around the environment in at least 4 directions relative to their forward-facing position.

- NFR-8: The graph **must** have transparent axes so the user is able to look through them and discern the dataset.
- NFR-9: Data points in the graph **must** be represented by spheres. These **should** be approximately a quarter graph unit in diameter.
- NFR-10: A subtle colour change **could** be used on the spheres along one axis, with the shift in colour representing distance along the axis.
- NFR-11: The data information overlay **should** detail the data point's values in all 3 dimensions.

A.4 Parallel Plane Requirements

- FN-39: Each new dimension in the dataset **must** be represented on a new plane.
- FN-40: The dimension in the z axis on each plane **should** be kept constant.
- FN-41: The user **must** be able to select and deselect a subset of data observations.
- FN-42: The user **should** be able to select and deselect multiple subsets of data observations.
- FN-43: Data observations across dimensions **must** be connected using a polyline. Its value on each plane is its intersection on that particular plane.

- NFR-12: When a subset of data is highlighted, it **must** appear in a highlighted colour. The rest of the dataset **should** be dimmed.

Appendix B

Privacy Challenges

The privacy challenges emanating from the collection of personal data were not considered to be within the scope of this project. Nevertheless, a short discussion continues below, characterising some of the recent research into big data and personal data.

"*The accumulation of personal data has an incremental adverse effect on privacy*" – Tene and Polonetsky (2012, p.251). Big data is increasingly used as a means for predictive analysis. Most recently, the insurance company Admiral was forced to abandon plans to analyse the social media history of car owners to set the price of insurance policies. Digital rights campaigners expressed concern about algorithms purposely biasing decisions "*based on race, gender, religion or sexuality*", and how this may encourage users to "*self-censor*" to affect future predictions made on big data sets (Killock, 2016). There are a vast number of papers and books written around the privacy challenges which big data presents (Thierer, 2015; Boyd and Crawford, 2011; Horvitz and Mulligan, 2015), particularly around health systems Mittelstadt and Floridi (2016). Self-tracking injects increasing amounts of personal data into big data systems, with specific concerns around technology enabling the recall of historical data (Nunan and Di Domenico, 2013) and anonymisation (Mittelstadt and Floridi, 2016). For instance, Fitbit's privacy policy[1] explicitly states that 'de-identified data' may be sold to interested audiences. However, research by Narayanan and Shmatikov (2008) shows that it is possible to combine large datasets and reidentify previously anonymised records.

A further privacy issue is related to data sharing and privacy policies. Given that self-tracking can involve deeply physiological data, it is essential that people know precisely which metrics are shared with companies. A 2016 study by the Future of Privacy Forum discovered that only 61% of the top health and fitness apps linked to a privacy policy from the app store listing page - 10% lower than apps present across all other categories. The study noted that although privacy policies are only a "*starting point for protecting individual's privacy, it is an important baseline standard around the world*" (Bates, 2016, p.3). Indeed, both iOS and Android have made ground in forcing people to actively participate

[1] https://www.fitbit.com/uk/privacy

with privacy, using granular permission controls and system dialogs prompting for access to specific sensors[2]. It follows that users would feel more control of their personal data with this method of participation. Nevertheless, Brandimarte et al. (2013) found that by providing people with more privacy controls, it encouraged users to publish more sensitive data. Paradoxically, this reduced the users level of privacy protection.

[2]https://developer.android.com/training/permissions/requesting.html

Appendix C

Project Plan

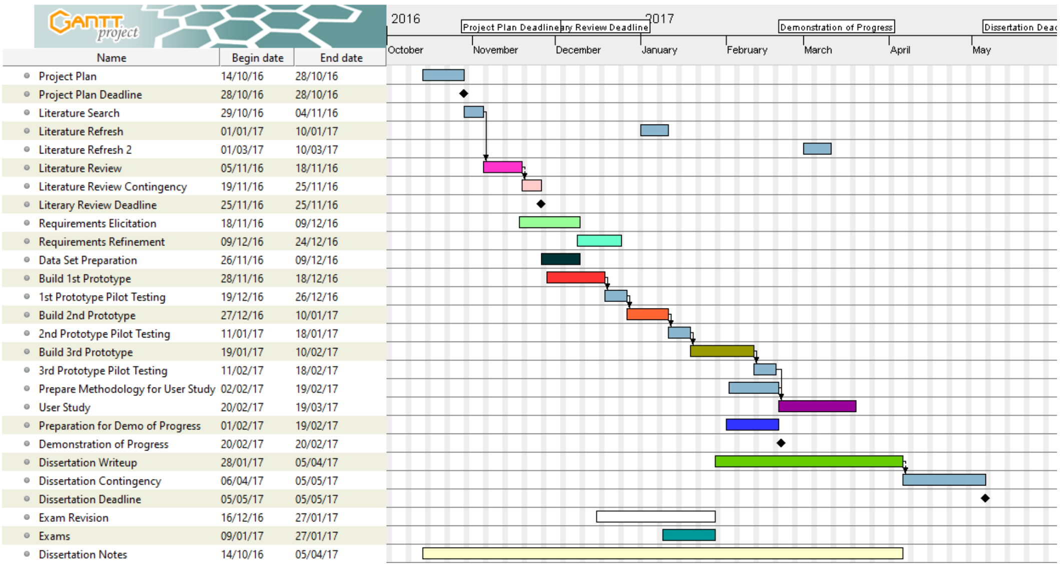


Figure C.1: Project Gantt Chart

A Gantt chart was prepared at the beginning of the project to structure the work we completed for this project in terms of the major deadlines. The overall structure was followed throughout, and provided a useful guide for visualising the dependencies between different tasks. Nevertheless, the project was typically about 1-2 weeks behind, with the exception of work relating to the Demonstration of Progress. The extensive amount of literature read for the Literature Review pushed us back a week in November, and the preparation of the format and materials for the User Study added another week later in March. In the future, we would add a task for *Results Analysis* after User Study, as this also took up a considerable amount of time.

Appendix D

Non-Linear Swipe-Based Movement Excerpt

Listing D.1: Movement via non-linear interpolation

```
// Allows for synchronousity with asynchronous child methods +
   deceleration
// Disable the gameobject once reaching it to enable raycaster
   to get out of the data point
private IEnumerator moveToWorldPositionCaller(float
   movementTime, Vector3 startPosition, Vector3 endPosition){
      enableColliding ();
      yield return interpolatePositions(movementTime,
         startPosition, endPosition, true);
      disableColliding ();
}

// Coroutine to yield new camera positions over time
private IEnumerator interpolatePositions(float movementTime,
   Vector3 startPosition, Vector3 endPosition, bool
   decelerationEnabled = false) {
      float i = 0.0f; //starting fraction
      float rate = 1.0f / movementTime; //rate at which i
         progresses from 0 to 1
      while (i < 1.0f) {
            //incrementally add to movement fraction: time
               of last frame rate production * rate of i
            i += Time.deltaTime * rate;
            // New camera positions based on chosen
               interpolation method
```

```
            if (decelerationEnabled) {
                yield return this.transform.parent.
                transform.position = Sinerp
                    (startPosition, endPosition, i);
            } else {
                yield return this.transform.parent.
                transform.position = Vector3.Lerp
                    (startPosition, endPosition, i);
            }
        }
}

// Enables non-linear interpolation. Sinerp is a curve with a
    small deceleration.
public static Vector3 Sinerp(Vector3 start, Vector3 end, float
    percentage){
        // Clamp to the range 0-1
        percentage = Mathf.Clamp01(percentage);
        // Follow sinerp instead of standard lerp for
            deceleration
        percentage = Mathf.Sin(percentage * Mathf.PI * 0.5f);
        return start + (end - start) * percentage;
}
```

All code is included in the zipped file submitted on Moodle. A *readme* is included which discusses the structure of the Unity code, and highlights the significant files we used.

Appendix E

Traditional 2D – Be The Data

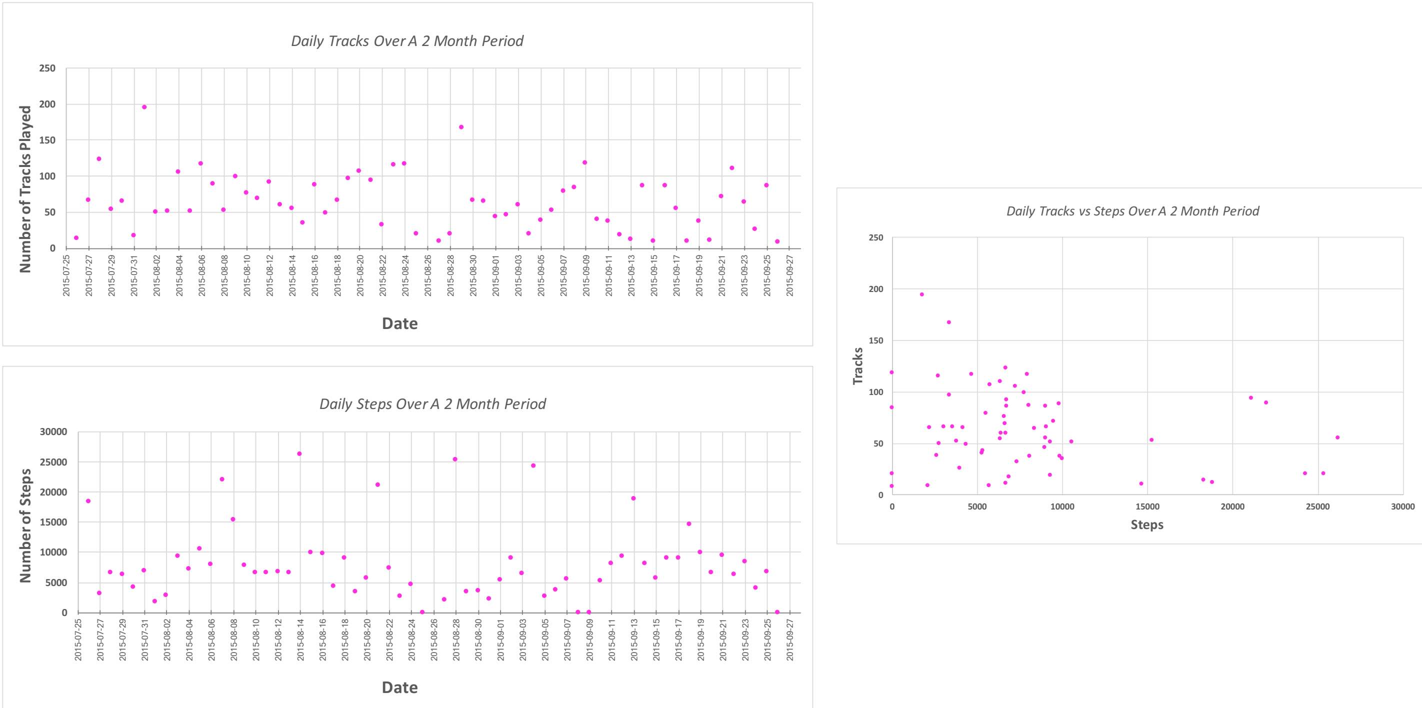


Figure E.1: Be The Data – 2D Training Visualisation

Figure E.2: Be The Data – 2D Experiment Visualisation

Appendix F

Traditional 2D – Parallel Planes

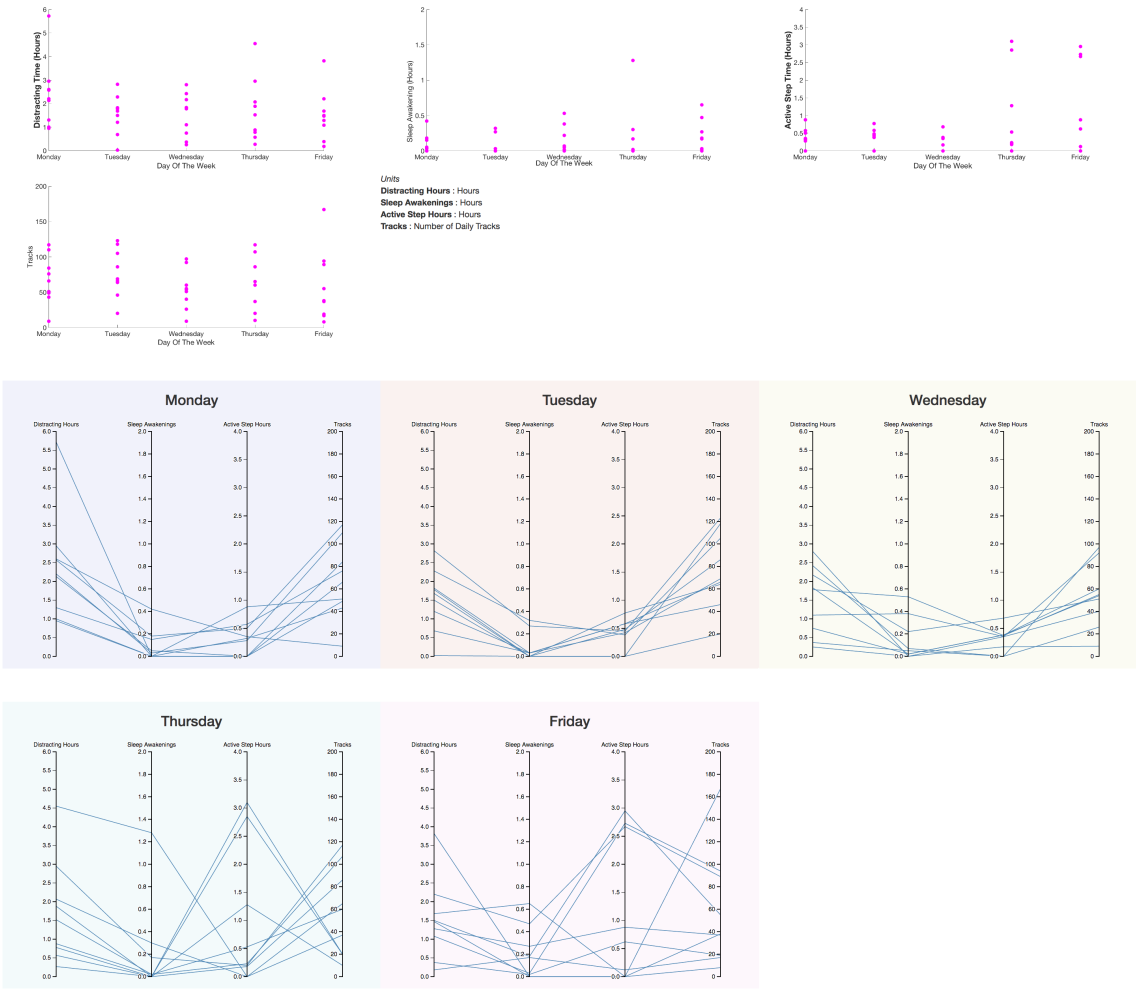


Figure F.1: Parallel Coordinates – 2D Training Visualisation

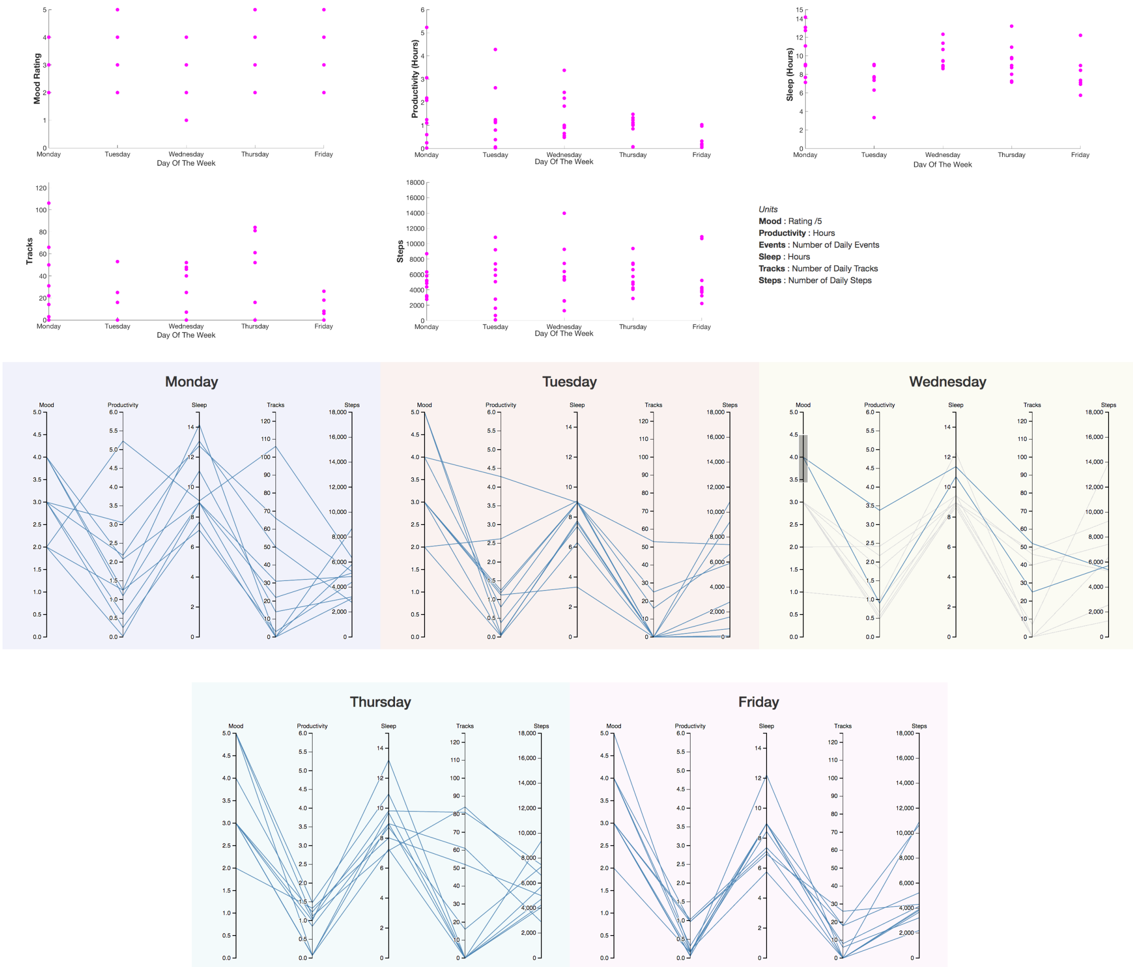


Figure F.2: Parallel Coordinates – 2D Experiment Visualisation (Includes brushing on Wednesday)

F.1 Matlab

Listing F.1: Code to create a scatter plot of a Parallel Plane for a selected dimension

```
disp(``Creating side-on perspective of Parallel Planes
   visualisation'');
table = readtable('2ddata.csv','Delimiter',',',
    'ReadVariableNames',false);

% Set to column of the dimension required to plot
dimension = 5;
% Set to 1 if keeping 0 valued data is important
% Note: This will not work with non-positive datasets
retain_zero = 1;


% Holds the final values to plot
x = zeros(height(table), 1);
y = zeros(height(table), 1);

% Start from the second row to skip the column headers
for i = 2:height(table)
    if table{i, 1} == ``Saturday'' || table{i, 1} == ``Sunday''
        continue % Ignore weekends
    else
        % Determine weekday and add to X vector
        if table{i, 1} == ``Monday''
            x(i) = 1;
        elseif table{i, 1} == ``Tuesday''
            x(i) = 2;
        elseif table{i, 1} == ``Wednesday''
            x(i) = 3;
        elseif table{i, 1} == ``Thursday''
            x(i) = 4;
        elseif table{i, 1} == ``Friday''
            x(i) = 5;
        end

        % Determine the value on that weekday and add to Y vector
        y(i) = str2double(table{i, dimension});

        % Special case for meaningful 0 data in a column
        % Overwrites 0s to -1s. Converts back to 0 after the loop
```

```
        if retain_zero == 1 && str2double(table{i, dimension})
           == 0
            y(i) = -1;
        end
    end
end

% Remove weekends in vectors
x(x==0) = [];
y(y==0) = [];

if retain_zero == 1
    y(y==-1) = 0;
end

% Now plot on a scatter graph
scatter(x,y);
```

Appendix G

13-Point Ethics Checklist

1. **Have you prepared a briefing script for volunteers?** Yes. The briefing script is included as part of the consent form which participants sign at the beginning. This makes clear what personal data we are collecting about them.

2. **Will the participants be using any non-standard hardware?** Yes. The Daydream headset is a VR headset which may result in nausea, motion sickness, or general discomfort. The consent form states this, and tells participants to remove the headset immediately if this occurs. They then have the opportunity to take a break, or leave the experiment if they do not wish to continue.

   We also make clear that participants must be aware of their surrounding environment while wearing the headset. We will also ensure that the participant is placed in an environment with an appropriate amount of physical space.

3. **Is there any intentional deception of the participants?** There is no intentional deception of participants.

   In the 2D visualisation we do not state that the study is about VR until the participant debrief. Nevertheless we assert that participants are unlikely to object or show unease when debriefed with this information.

4. **How will participants voluntarily give consent?**

   We will use our consent form to record the participant's signature. If appropriate, this enables the results to be used beyond this study.

5. **Will the participants be exposed to any risks greater than those encountered in their normal work life?** No.

6. **Are you offering any incentive to the participants?** No.

7. **Are any of your participants under the age of 16?** No.

8. **Do any of your participants have an impairment that will limit their understanding or communication?** No.
9. **Are you in a position of authority or influence over any of your participants?** No.
10. **Will the participants be informed that they could withdraw at any time?** Yes, this is included on the consent form and will be repeated verbally.
11. **Will the participants be informed of your contact details?** Yes, this is included on the consent form.
12. **Will participants be de-briefed?** Yes. An overview of the aims of the experiment will be provided and how their involvement supports these aims.
13. **Will the data collected from the participants be stored in an anonymous form?** Yes.

Name: Patrick Millais

Supervisor: Dr Simon Jones

Project Title: Visualisation and Exploration of Personal Data in Virtual Reality

Date: 02/03/2017

Appendix H

Participant Consent Forms

H.1 Traditional 2D Consent Form

The purpose of this experiment is to capture user insight resulting from interaction with various data visualisations. In turn, this will enable us to measure the effectiveness of specific visualisations and tools at supporting users to explore and arrive at insights. Please note that this experiment is not testing your personal intelligence! We value your honest feedback.

The experiment consists of the following parts: *Participant Brief, Training Stage, Main Stage, Participant Debrief*

The following participant data will be recorded at the beginning of the experiment: *Age, Gender, Education Level, GCSE/A Level Maths.*
All data will be stored securely, anonymously and only accessible by the main researcher and supervisor. Audio will also be recorded during the training and main stages.

If you are interested in the results of this study, or would like to get in touch with the researcher, you can contact Patrick Millais (pm515bath.ac.uk).

Your participation in this study is completely voluntary and you are free to withdraw at any time, without giving a reason and without any consequences.

You now have the opportunity to ask questions and discuss the experiment further with the researcher.

Your signature indicates that you have read the above, and that you have received enough information about this study and consent to participating. Your participation is voluntary and can be withdrawn at any time.

Signed ______________________________ Date: __________

Researcher: Patrick Millais
Project Supervisor: Dr Simon Jones

H.2 VR Consent Form

The purpose of this experiment is to capture user insight resulting from interaction with a data visualisation in Virtual Reality. In turn, this will enable us to measure the effectiveness of specific visualisations and tools at supporting users to explore and arrive at insights. Please note that this experiment is not testing your personal intelligence! We value your honest feedback.

The experiment consists of the following parts: *Participant Brief, Training Stage, Main Stage, Participant Debrief*

The following participant data will be recorded at the beginning of the experiment: *Age, Gender, Education Level, GCSE/A Level Maths.*
All data will be stored securely, anonymously and only accessible by the main researcher and supervisor. Audio will also be recorded during the training and main stages.

If you are interested in the results of this study, or would like to get in touch with the researcher, you can contact Patrick Millais (pm515bath.ac.uk).

Your participation in this study is completely voluntary and you are free to withdraw at any time, without giving a reason and without any consequences.

This study uses a Virtual Reality headset and controller to display data visualisations. The use of this hardware may result in nausea, motion sickness or general discomfort. If this occurs, remove the headset immediately and take a break. If you do not feel comfortable continuing, please inform the researcher and the experiment will be stopped. Please also be aware of the surrounding environment when in VR.

You now have the opportunity to ask questions and discuss the experiment further with the researcher.

Your signature indicates that you have read the above, and that you have received enough information about this study and consent to participating. Your participation is voluntary and can be withdrawn at any time.

Signed ______________________________ Date: __________

Researcher: Patrick Millais
Project Supervisor: Dr Simon Jones

Appendix I

Questionnaires

I.1 NASA Task Load Index (NASA-TLX)

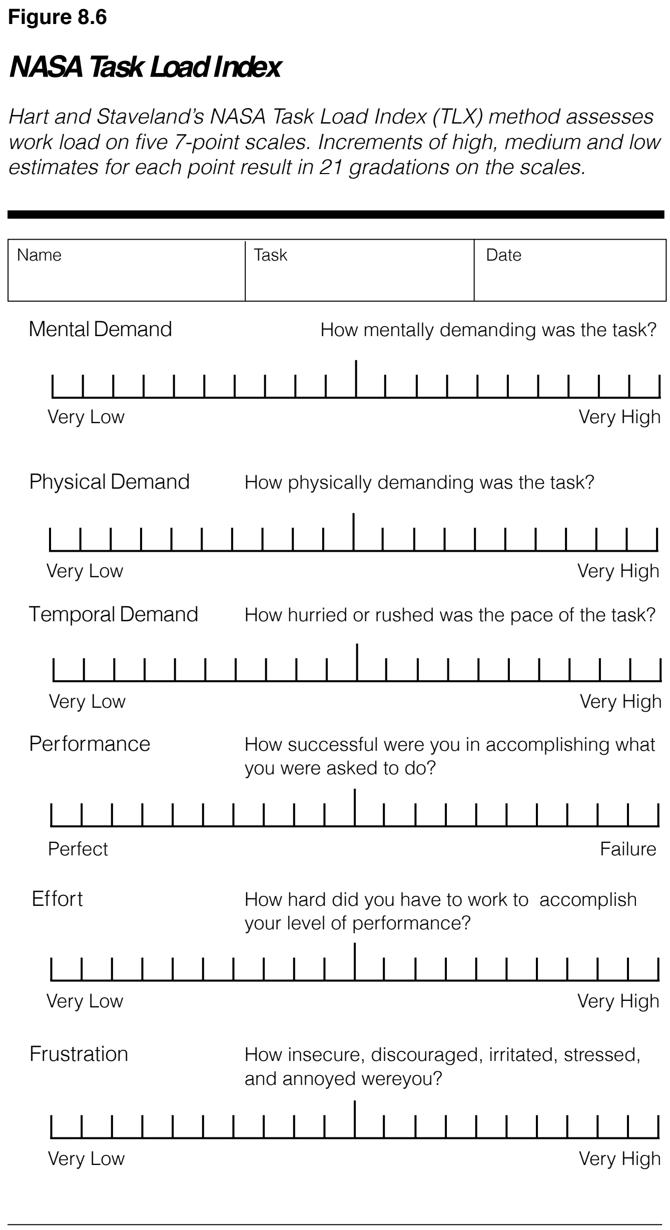


Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure I.1: NASA-TLX Scales

I.1.1 NASA-TLX Definitions

Mental Demand *(Low/High)*
How much mental and perceptual activity was required (for example, thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, forgiving or exacting?

Physical Demand *(Low/High)*
How much physical activity was required (for example, pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

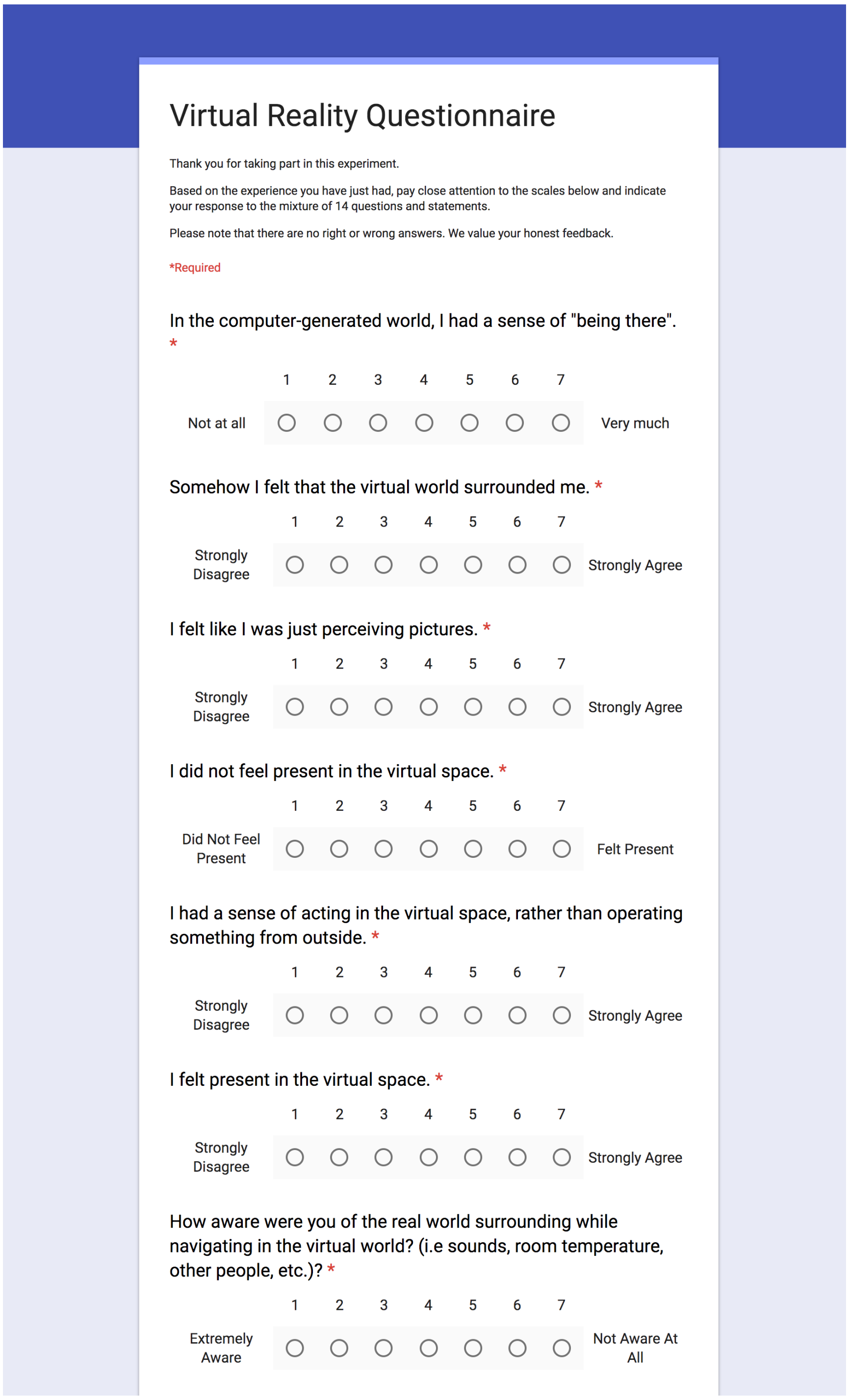
Temporal Demand *(Low/High)*
How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Performance *(Good/Poor)*
How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Effort *(Low/High)*
How hard did you have to work (mentally and physically) to accomplish your level of performance?

Frustration Level *(Low/High)*
How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

I.2 IPQ – Presence Questionnaire



Virtual Reality Questionnaire

Thank you for taking part in this experiment.

Based on the experience you have just had, pay close attention to the scales below and indicate your response to the mixture of 14 questions and statements.

Please note that there are no right or wrong answers. We value your honest feedback.

*Required

In the computer-generated world, I had a sense of "being there". *

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very much

Somehow I felt that the virtual world surrounded me. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

I felt like I was just perceiving pictures. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

I did not feel present in the virtual space. *

	1	2	3	4	5	6	7	
Did Not Feel Present	○	○	○	○	○	○	○	Felt Present

I had a sense of acting in the virtual space, rather than operating something from outside. *

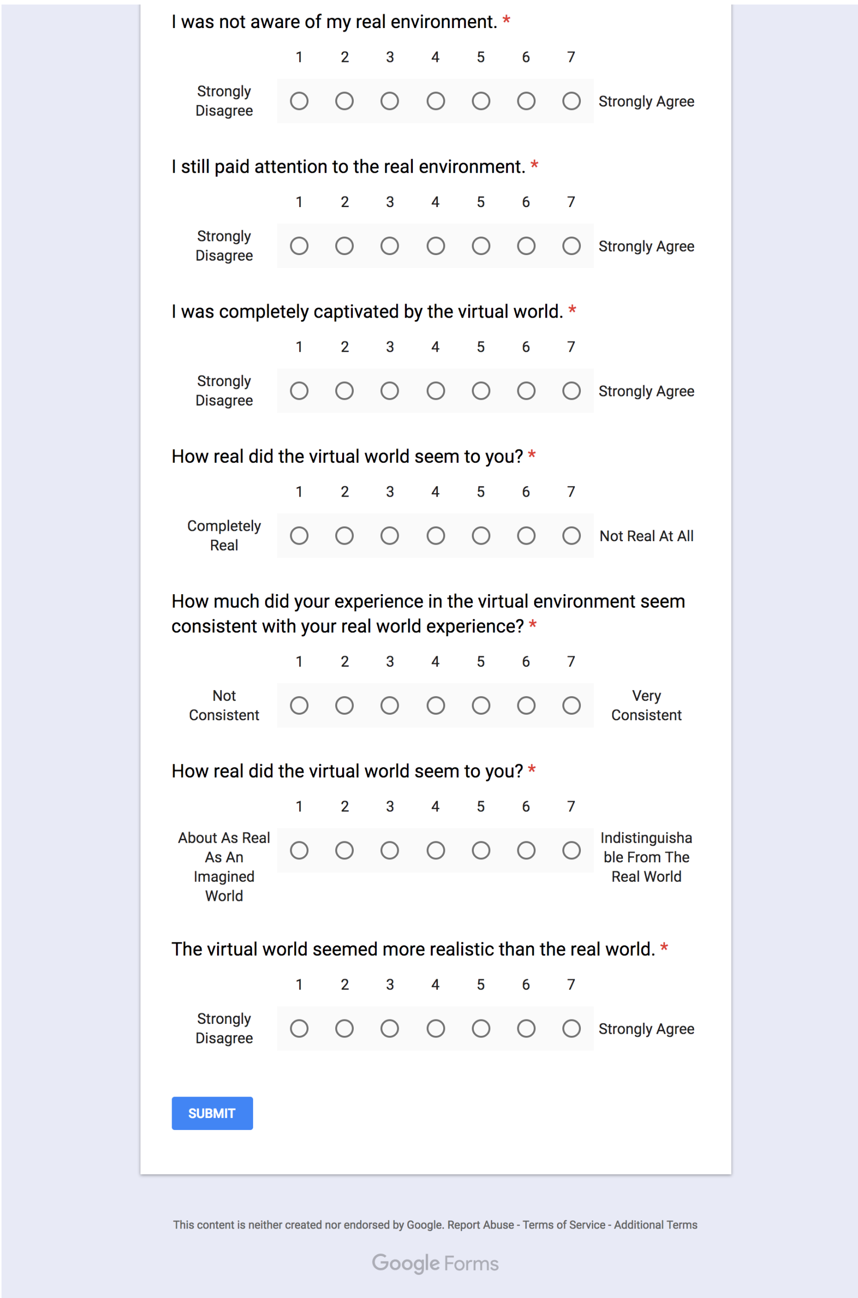
	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

I felt present in the virtual space. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

How aware were you of the real world surrounding while navigating in the virtual world? (i.e sounds, room temperature, other people, etc.)? *

	1	2	3	4	5	6	7	
Extremely Aware	○	○	○	○	○	○	○	Not Aware At All



I was not aware of my real environment. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

I still paid attention to the real environment. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

I was completely captivated by the virtual world. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

How real did the virtual world seem to you? *

	1	2	3	4	5	6	7	
Completely Real	○	○	○	○	○	○	○	Not Real At All

How much did your experience in the virtual environment seem consistent with your real world experience? *

	1	2	3	4	5	6	7	
Not Consistent	○	○	○	○	○	○	○	Very Consistent

How real did the virtual world seem to you? *

	1	2	3	4	5	6	7	
About As Real As An Imagined World	○	○	○	○	○	○	○	Indistinguishable From The Real World

The virtual world seemed more realistic than the real world. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

SUBMIT

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Additional Terms

Google Forms

Figure I.2: IPQ – Presence Questionnaire

Appendix J

Example Data Analysis

During our empirical evaluation we analysed transcribed and analysed 16 recordings of participants interacting with 2D and VR data visualisations. The recordings, transcripts, and full analysis for all 16 participants are included in the zipped folder submitted on Moodle.

A short excerpt from one participant's data analysis is submitted here as an example. Table J.1 was for Participant 16 who interacted with the VR Parallel Planes experiment. Their initial exploratory questions were:

- Correlation between Mood vs. Steps
- Correlation between Productivity vs. Sleep
- Correlation between Sleep and all others

This table forms a small section of the number of insights which this participant gave. Their full results, alongside analysis for all other participants, are included in the submitted zip file on Moodle.

Table J.1: Participant 16 Analysis Excerpt

Observation	Breadth or Depth	Directed or Unexpected	Hypotheses	Domain Value	Correctness	Category
General pattern is down up down up	Breadth	Unexpected		1	Correct	Pattern
So I've selected an outlier with a high number of tracks, with a middling number of steps, quite a high sleep, really high productivity, and middling mood. This person listened to a lot of music and was quite productive on that day.	Depth	Unexpected		2	Correct	Detail
Another day with high sleep, but productivity is low and their music is low. So that's completely opposite. So you either sleep or listen to loads of music to be productive.	Depth	Unexpected	Yes	3	Correct	Pattern
Productivity to sleep is strongly correlated	Breadth	Directed		2	Correct	Pattern
The tracks just seems quite a random thing in relation to all the others	Breadth	Unexpected		2	Correct	Pattern
I'm now looking at really low tracks, just to see if I can find anything. Because there's some which are right at the bottom and that has some of the least step days actually. Some of the lowest sleeps days as well actually	Depth	Unexpected	Yes	4	Correct	Group
Less music = less sleep, which is a bit weird because you probably want to listen to music when you are awake more	Breadth	Directed		3	Correct	Pattern

Appendix K

Evaluation Results

K.1 Presence Results

Table K.1: Raw Igroup Presence Questionnaire (IPQ) Results

ID	Experiment	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	T	T*
1	VR Be The Data	7	6	2	7	5	7	6	6	1	6	6	1	1	1	62	68
5	VR Parallel Planes	7	7	1	6	7	6	4	6	2	5	4	6	5	2	68	78
7	VR Be The Data	6	7	1	7	3	7	6	6	1	6	2	3	3	1	59	75
10	VR Be The Data	6	5	3	6	3	5	3	5	4	5	3	3	3	2	56	60
11	VR Be The Data	5	6	2	6	6	6	7	7	2	5	5	3	3	1	64	70
12	VR Parallel Planes	5	4	2	6	4	6	6	5	3	3	6	2	1	1	54	56
13	VR Parallel Planes	6	7	2	6	6	6	4	4	2	3	6	4	4	1	61	65
16	VR Parallel Planes	5	3	3	6	4	6	5	6	4	5	4	5	2	3	61	63

T is the composite score for the 14 questions.
T* is the composite score which accounts for reverse coded questions.

K.2 Task Workload Results

K.2.1 Raw Results

Table K.2: Raw NASA-TLX Results

		TLX Subscale						
ID	Experiment	Mental	Physical	Temporal	Performance	Effort	Frustration	**Overall**
1	VR Be The Data	40	70	5	10	45	30	33.3
2	2D Parallel Planes	70	15	20	80	55	20	43.3
3	2D Parallel Planes	65	5	25	45	50	25	35.8
4	2D Be The Data	55	5	5	30	30	5	21.7
5	VR Parallel Planes	25	5	15	20	25	5	15.8
6	2D Be The Data	70	5	15	70	35	65	43.3
7	VR Be The Data	90	25	5	25	85	15	40.8
8	2D Be The Data	50	5	15	60	55	55	40.0
9	2D Parallel Planes	70	5	10	60	60	55	43.3
10	VR Be The Data	45	25	30	30	60	25	35.8
11	VR Be The Data	50	40	55	30	45	20	40.0
12	VR Parallel Planes	50	60	35	40	45	45	45.8
13	VR Parallel Planes	35	20	20	35	50	55	35.8
14	2D Parallel Planes	25	20	25	45	40	10	27.5
15	2D Be The Data	35	5	25	45	50	60	36.7
16	VR Parallel Planes	95	25	50	40	80	60	58.3

K.2.2 Normality Tests

Table K.3: TLX Subscale Normality Tests

Variable	**W**	**df**	**Sig.**
Task Work Load Index			
2D - Be The Data	0.87	4	0.296
2D - Parallel Planes	0.861	4	0.264
VR - Be The Data	0.912	4	0.492
VR - Parallel Planes	0.985	4	0.933
Effort			
2D - Be The Data	0.911	4	0.488
2D - Parallel Planes	0.971	4	0.85
VR - Be The Data	0.841	4	0.199
VR - Parallel Planes	0.963	4	0.797
Frustration			
2D - Be The Data	0.763	4	0.051
2D - Parallel Planes	0.897	4	0.414
VR - Be The Data	0.993	4	0.972
VR - Parallel Planes	0.833	4	0.177
Mental			
2D - Be The Data	0.991	4	0.962
2D - Parallel Planes	0.708	4	0.014
VR - Be The Data	0.79	4	0.085
VR - Parallel Planes	0.893	4	0.395
Performance			
2D - Be The Data	0.979	4	0.894
2D - Parallel Planes	0.854	4	0.241
VR - Be The Data	0.791	4	0.086
VR - Parallel Planes	0.791	4	0.086
Physical			
2D - Be The Data	-	-	-
2D - Parallel Planes	0.849	4	0.224
VR - Be The Data	0.827	4	0.161
VR - Parallel Planes	0.918	4	0.528
Temporal			
2D - Be The Data	0.945	4	0.683
2D - Parallel Planes	0.827	4	0.161
VR - Be The Data	0.863	4	0.272
VR - Parallel Planes	0.94	4	0.653

K.2.3 Levene's Test of Equality

	F	df1	df2	Sig.
Overall TLX	2.208	3	12	0.14
Effort	0.714	3	12	0.562
Frustration	1.445	3	12	0.279
MentalDemand	0.575	3	12	0.642
Performance	1.222	3	12	0.344
PhysicalDemand	2.739	3	12	0.09
TemporalDemand	3.606	3	12	0.046

Table K.4: TLX Test of Equality

K.2.4 Overall Task Load Index Mean & Standard Deviation

Medium		**Mean**	**Std. Deviation**	N
2D	Be The Data	35.4167	9.56217	4
	Parallel Planes	37.5000	7.54615	4
	Total	*36.4583*	*8.05179*	*8*
VR	Be The Data	37.5000	3.53553	4
	Parallel Planes	38.9583	17.95538	4
	Total	*38.2292*	*12.00560*	*8*
Total	Be The Data	36.4583	6.76637	8
	Parallel Planes	38.2292	12.77427	8
	Total	*37.3437*	*9.91734*	*16*

Table K.5: Overall Task Load Index Mean & Standard Deviation

K.2.5 Subscale Means of Task Workload

		Mental	**Physical**	**Temporal**	**Performance**	**Effort**	**Frustration**	**Overall**
2D	**BTD The Data**	52.5	5	15	51.25	42.5	46.25	35.42
	Parallel Planes	57.5	11.25	20	57.5	51.25	27.5	37.5
	Combined	55	8.13	17.5	54.38	46.88	36.88	36.46
VR	**Be The Data**	56.25	40	23.75	23.75	58.75	22.5	37.5
	Parallel Planes	51.25	27.5	30	33.75	50	41.25	38.96
	Combined	53.75	33.75	26.88	28.75	54.38	31.88	38.23

Table K.6: Analysis of means on individual task workload subscales

K.3 Insight Results

K.3.1 Normality Tests

Shapiro-Wilk		**W**	**dF**	**p**
	2D Be The Data	0.848	4	0.22
Number of Insights	2D Parallel Planes	0.993	4	0.972
EH2	VR Be The Data	0.945	4	0.683
	VR Parallel Planes	0.982	4	0.911
	2D Be The Data	0.801	4	0.103
Number of Correct Insights	2D Parallel Planes	0.96	4	0.78
EH3	VR Be The Data	0.773	4	0.062
	VR Parallel Planes	0.827	4	0.161
	2D Be The Data	0.863	4	0.272
Number of Incorrect Insights	2D Parallel Planes	0.827	4	0.161
EH3	VR Be The Data	0.729	4	0.024
	VR Parallel Planes	0.63	4	0.001
Number of Unexpected	2D	0.732	8	0.005
EH4	VR	0.816	8	0.042
Depth Insights	2D Be The Data	0.993	4	0.972
EH5	2D Parallel Planes	0.863	4	0.272
	VR Be The Data	0.895	4	0.406
	VR Parallel Planes	0.927	4	0.577

Table K.7: Shapiro-Wilk Tests for Assumption of Normality

K.3.2 Levene's Test of Equality

	F	df1	df2	Sig.
Number of Insights	2.546	3	12	0.105
Correct Insights	2.303	3	12	0.129
Incorrect Insights	1.372	3	12	0.298
Unexpected Insights	5.661	3	12	0.012
Hypotheses	1.418	3	12	0.286
Depth Insights	0.535	3	12	0.667

Table K.8: Levene's Test of Equality for dependent insight variables

K.3.3 Analysis of Insights

	2D Be The Data				VR Be The Data			
	ID 4	*ID 6*	*ID 8*	*ID 15*	*ID 1*	*ID 7*	*ID 10*	*ID 11*
Number of Insights	9	11	9	14	12	13	13	14
Breadth Insights	7	8	9	13	9	11	10	9
Depth Insights	2	3	0	1	3	2	3	5
Directed Insights	9	11	9	13	11	13	13	14
Unexpected Insights	0	0	0	1	1	0	0	0
Hypotheses	0	0	0	2	2	0	0	0
Domain Value = 1	2	4	4	1	4	6	6	6
Domain Value = 2	5	7	5	9	4	7	7	8
Domain Value = 3	2	0	0	4	2	0	0	0
Domain Value = 4	0	0	0	0	2	0	0	0
Mean Domain Value	3.0	9.0	7.0	11.0	2.2	1.5	1.5	1.6
Correctness = Correct	6	2	1	2	9	9	8	14
Correctness = Neutral	0	0	1	1	3	3	4	0
Correctness = Incorrect	4	5	4	6	0	1	1	0
Category = Overview	2	3	3	8	4	5	4	4
Category = Pattern	1	1	2	0	5	6	7	6
Category = Group	0	2	0	0	1	1	1	1
Category = Detail	0	2	0	0	2	1	1	3

Table K.9: Insight Matrix – Part 1/2

	2D Parallel Planes				VR Parallel Planes			
	ID 2	*ID 3*	*ID 9*	*ID 14*	*ID 5*	*ID 12*	*ID 13*	*ID 16*
Number of Insights	27	25	26	28	27	21	15	23
Breadth Insights	13	9	10	16	18	14	7	11
Depth Insights	14	16	16	12	9	7	8	12
Directed Insights	23	22	20	11	15	12	7	10
Unexpected Insights	4	3	6	17	12	9	8	13
Hypotheses	7	2	4	6	4	8	4	5
Domain Value = 1	7	6	7	5	3	3	0	1
Domain Value = 2	13	17	16	17	19	10	10	16
Domain Value = 3	7	1	2	4	5	2	4	5
Domain Value = 4	0	1	1	2	0	6	1	1
Mean Domain Value	2.0	1.9	1.9	2.1	2.1	2.5	2.4	2.3
Correctness = Correct	11	24	15	21	21	18	12	21
Correctness = Neutral	14	0	7	6	5	3	2	1
Correctness = Incorrect	2	1	4	1	1	0	1	1
Category = Overview	9	4	5	11	2	4	1	0
Category = Pattern	13	9	8	7	16	10	7	15
Category = Group	2	1	0	0	5	2	1	3
Category = Detail	3	11	13	10	4	5	6	5

Table K.10: Insight Matrix – Part 2/2